

# Real-time dynamics acquisition from irregular samples – with application to anesthesia evaluation

Charles K. Chui

*Department of Statistics, Stanford University, Stanford, CA 94305, USA (ckchui@stanford.edu)*

Yu-Ting Lin

*Department of Anesthesiology, Shin Kong Wu Ho-Su Memorial Hospital, Taipei, Taiwan; Graduate Institute of Biomedical Electronics and Bioinformatics, National Taiwan University, Taipei, Taiwan (linyuting@hotmail.com.tw)*

Hau-tieng Wu

*Department of Mathematics, Stanford University, Stanford, CA 94305, USA (hawwu@stanford.edu)*

---

## Abstract

Although digital representations of information sources are ubiquitous in our daily life, almost all digital information sources are regular samples (or obtained by uniform sampling) of some continuous function representations. However, there are many important events for which only irregular data samples are available, including trading data of the financial market and various clinical data, such as the respiration signals hidden in ECG measurements. For such digital information sources, the only available effective smooth function interpolation scheme for digital-to-analog (D/A) conversion algorithms are mainly for off-line applications. Hence, in order to adapt the powerful continuous-function mathematical approaches for real-time applications, it is necessary to introduce an effective D/A conversion scheme as well as to modify the desired continuous-function mathematical method for on-line implementation. The powerful signal processing tool to be discussed in this paper is the synchrosqueezing transform (SST), which requires computation of the continuous wavelet transform (CWT), as well as its derivative, of the analog signal of interest. An important application of this transform is to extract information, such as the underlying dynamics, hidden in the signal representation. The first objective of this paper is to introduce a unified approach to remove the two main obstacles for adapting the SST approach to irregular data samples in order to allow online computation. Firstly, for D/A conversion, a real-time algorithm, based on spline functions of arbitrarily desired order, is proposed to interpolate the irregular data samples, while preserving all polynomials of the same spline order, with assured maximum order of approximation. Secondly, for real-time dynamic information extraction from an oscillatory signal via SST, a family of vanishing-moment and minimum-supported spline-wavelets (to be called VM wavelets) are introduced for on-line computation of the CWT and its derivative. The second objective of this paper is to apply the proposed real-time algorithm and VM wavelets to clinical applications, particularly to the study of the “anesthetic depth” of a patient during surgery, with emphasis on analyzing two dynamic quantities: the “instantaneous frequencies” and the “non-rhythmic to rhythmic ratios” of the patient’s respiration, based

on a one-lead electrocardiogram (ECG) signal. Indeed, the “R-peaks” of the ECG signal, which constitute a waveform landmark for clinical evaluation, are non-uniform samples of the respiratory signal. It is envisioned that the proposed algorithm and VM wavelets should enable real-time monitoring of “anesthetic depth”, during surgery, from the respiration signal via ECG measurement.

---

## 1. Introduction

Digital representation of time series, signals, images, videos, and other information sources is ubiquitous in our daily life. In almost all application areas, digital representation is achieved by sampling or discretization of a continuous function. For efficiency, uniform sampling is used to give regular data samples. This means that sampling is carried out at uniformly (or equally) spaced points of the time or spatial domain. On the other hand, there are many important events for which only irregular data samples are available. These include trading data of the financial market, package data traffic [31, 36], measurement of star luminosity [57] (due to weather condition or day/night difference), and clinical data, such as respiration signals (hidden in ECG measurements) [48]. Other irregular data samples are introduced by non-uniform sampling to meet certain requirements, such as speeding up MRI (magnetic resonance image) acquisition [38] and increasing performance of radar systems to avoid jamming [9] and to minimize aliasing [1, 6].

In any case, for such application areas as mentioned above, we need effective methods to process non-uniform data samples in order to obtain the information of interest, and more importantly, to be able to apply the powerful tools that apply only to analog data representations. In other words, we need an effective D/A (digital to analog) conversion tool. Of course there are various D/A methods for non-uniform data samples introduced for different purposes that are available in the literature; for example, finding a proper building block under certain reconstruction requirements [24, 5, 14, 27], applying non-parametric regression to irregular data fitting [28], and iterative reconstruction algorithms or projection onto suitable convex sets [44, 32], just to name a few. However, all of these and other approaches, including recovery of the continuous-time function from its Fourier transform, estimated directly from the irregular sample [41, 47], are not suitable for D/A conversion in a real-time or online system. Examples include real-time warning of a network under attack from non-uniformly sampled packet data traffic, and for the clinical application to be studied in this paper, real-time warning if a patient is entering a critical condition.

A powerful signal processing tool that applies only to continuous-time signals is the synchrosqueezed continuous wavelet transform (SST), which requires knowledge of the continuous wavelet transform (CWT) of the signal [22]. An important application of this transform is to extract information, such as the underlying dynamics, hidden in the signal representation [22, 12]. The main benefit of choosing CWT is its polynomial cancelation property, which allows us to handle the trend in real time. However, there are two obstacles for real-time application of the SST. Firstly, an effective algorithm for D/A conversion of non-uniform samples is needed. Although a partial off-line solution to meet this challenge is proposed in [55, 58] based on the synchrosqueezed short-time Fourier transform (STFT) of non-uniform data samples, there are various short-comings of this direct approach. In particular, due to the lack of sufficient vanishing moments,

the possible trend that commonly exists in the signal cannot be eliminated. On the other hand, in order to achieve precise spectral information, the analysis wavelet for the CWT is assumed to have compactly supported Fourier transform [22], and consequently the SST cannot be realized in real-time. Of course an obvious suggestion is to replace the analysis wavelet by a compactly supported wavelet. However, the obstacle remains in that for real-time realization, the CWT has to be taken on a bounded time-interval, and that a high order of vanishing moments of the analysis wavelet is required for reasonably good performance of the SST, which in turn requires larger support of the wavelet.

One of the main objectives of this paper is to introduce a unified approach to remove, or at least minimize, the two obstacles mentioned above. Firstly, for D/A conversion, a real-time algorithm, based on spline functions of arbitrarily desired order, is proposed to interpolate the non-uniform data, while preserving all polynomials of the same order, with assured maximum order of approximation. Secondly, for real-time dynamic information extraction from an oscillatory signal via SST, we introduce a family of analysis wavelets with minimum support in the time-domain, again in terms of B-splines of arbitrarily desired order, for both the interior and the boundary (i.e. at the end-points) of any bounded interval, with arbitrarily desired order of vanishing moments. We coin these analysis wavelets as VM wavelets (for maximum vanishing moments and minimum support). We give explicit formulas of the VM wavelets on arbitrarily desired knot sequences, and prove that for the VM wavelets  $\psi_{m,n}$ , in terms of the  $m^{th}$  Cardinal B-splines (i.e. on integer knots) and with  $n^{th}$  order vanishing moments, a suitable scaling of the centered wavelets  $\psi_{m,n}$  is asymptotically the same as the  $n^{th}$  derivative of the Gaussian function, as  $m + n$  tends to infinity. With a VM wavelet as the mother wavelet of the CWT, computation of its derivative can be eliminated by replacing the CWT with a companion VM wavelet.

To demonstrate the usefulness of the proposed real-time algorithm and VM wavelets is applied to clinical applications, in the study of the anesthetic depth of a patient during surgery. More precisely, we will focus on analyzing two dynamic quantities: the instantaneous frequencies and the non-rhythmic to rhythmic ratios of the patient's respiration, by analyzing solely the one-lead electrocardiogram (ECG) signal. Indeed, the R-peaks of the ECG signal, that constitute a waveform landmark, are non-uniform samples of the respiratory signal. The proposed algorithm and VM wavelets enable us to study the possibility of real-time monitoring the anesthetic depth during surgery, from the respiration signal via ECG measurement.

This paper is organized as follows. In Section 2, the theoretical background of B-splines on arbitrary knot sequences, quasi-interpolation operation, complete local spline interpolation, and the blending operator are reviewed; a real-time algorithm for spline interpolation of non-uniform data samples is formulated; and the order of maximum order of approximation by the real-time interpolation algorithm is derived. The notion of VM wavelets is introduced in Section 3, where both interior and boundary wavelets with minimum support are derived, with explicit formulas, even on arbitrary knot sequences. Furthermore, the performance of the VM wavelets is analyzed in terms of the side-lobe/main lobe ratios of their powers spectra; and their asymptotic behavior, as compared with derivatives of the Gaussian function, is also derived in Section 3. In Section 4, the synchrosqueezing transform (SST) is reviewed and the notion of time-varying power spectrum (tvPS) is discussed, with the goal of extracting dynamic information from an

oscillating signal. Section 5 is devoted to the study of clinical applications of the proposed algorithm. In particular, the *blending ECG derived respiration (EDR) algorithm* is proposed to extract the respiratory dynamics from the single ECG signal in real-time and to study the relationship between anesthetic depth and respiratory dynamical features extracted from analyzing the morphology of the single lead ECG signal.

**Notation:** In this paper, the Fourier transform of a function  $h$  in  $L^2(\mathbb{R})$  is defined as  $\hat{h}(\xi) := \int_{-\infty}^{\infty} h(x)e^{-ix\xi}dx$ .

## 2. Optimal real-time spline interpolation

Let  $\{x_n\}$ ,  $n = 0, 1, \dots$ , be a time series contaminated with noise, where

$$x_n = g(t_n) + \epsilon_n, \quad (1)$$

is the non-uniform measurement of an underlying signal  $g(t)$  at the time positions  $t = t_n$ . An important problem in signal processing is to convert  $\{x_n\}$  on-line to an analog signal  $\tilde{g}(t)$ , where  $t \geq t_0$ , under certain requirements and having certain desirable properties. Here,  $\{\epsilon_n\}$  is a sequence of additive noise, including instrumental noise, electoral contact noise, baseline drift/motion noise, and statistical random noise. Since there has been a lot of study on detection and removal of deterministic noise, it is safe to assume that  $\{\epsilon_n\}$  is a random noise process.

For (statistical) white noise  $\{\epsilon_n\}$  with variance  $\sigma^2 > 0$ , the most efficient real-time algorithm to optimal recover  $\{g(t_n)\}$  from  $\{x_n\}$  is Kalman filtering. In a nutshell, the Kalman noise-removal filter is a prediction-correction algorithm that corrects the prediction  $\hat{x}_{n,n-1}$  to yield the optimal estimation  $\hat{x}_n = \hat{x}_{n,n}$  of  $x_n$ , in the sense that  $\{\hat{x}_k\}_{k=0}^n$  is the weighted least-squares fit of the measurement  $\{x_k\}_{k=0}^n$ , with  $1/\sigma^2$  as the weight. The interested reader is referred to [16], where white noise is also extended to color noise.

In this paper, we are only concerned with real-time conversion of a “clean” digital signal to spline function representation. In other words, we will assume that  $\hat{x}_n = x_n = g(t_n)$ , for all  $n = 0, 1, \dots$ . Let us first pose the problem and introduce some notations. The digital samples  $g(t_n)$  are taken at

$$\underline{t} : a = t_0 < t_1 < t_2 < \dots, \quad (2)$$

where  $\{t_k\}$  may be irregular (or non-uniform), meaning that we allow  $t_{k+1} - t_k \neq t_{j+1} - t_j$  for  $k \neq j$ . Let  $m \geq 3$  be any desired integer and  $\Pi_{m-1}$  denote the space of all polynomials of degree less than  $m$ .

**Definition** [Spline space] For any sequence

$$\underline{s} : s_{-m+1} = s_{-m+2} = \dots = a = s_0 < s_1 < s_2 < \dots, \quad (3)$$

we use the notation  $S_{\underline{s},m} := S_{\underline{s},m}[a, \infty)$  to denote the spline space of order  $m$  with knot sequence  $\underline{s}$ ; that is,  $f \in S_{\underline{s},m}$  if and only if  $f \in C^{m-2}[0, \infty)$  and  $f|_{[s_k, s_{k+1}]} \in \Pi_{m-1}$  for all  $k = 0, 1, \dots$ .

To formulate a locally supported basis of  $S_{\underline{s},m}$ , we first introduce the notion of *truncated powers*

$$x_+^{m-1} := (\max\{0, x\})^{m-1}.$$

Then it is clear that the truncated power functions  $(s_k - t)_+^{m-1}$ ,  $k = 0, 1, \dots$ , are in  $S_{\underline{s},m}$ . Since  $(x_k - t)_+^{m-1}$ ,  $k \geq 1$ , have global support, we apply the  $m$ -th order divided differences to change them to locally supported functions. In general, the *divided differences* are defined by

$$[u, \dots, u]f := \frac{f^{(l)}(u)}{l!}$$

if there are  $l + 1$  entries in  $[u, \dots, u]$ , and

$$[u_0, \dots, u_n]f := \frac{[u_1, \dots, u_n]f - [u_0, \dots, u_{n-1}]f}{u_n - u_0}$$

if  $u_0 \leq u_1 \leq \dots \leq u_n$  and  $u_n > u_0$ .

**Definition** [B-spline [24]] Fix  $m > 0$  and consider the knot sequence  $\underline{s}$  of the spline space  $S_{\underline{s},m}$ . The *normalized B-splines* are defined by applying the divided difference operation to the truncated power  $(s_k - t)_+^{m-1}$ , namely,

$$N_{m,k}(t) = N_{\underline{s},m,k}(t) = (s_{m+k} - s_k)[s_k, \dots, s_{m+k}](\cdot - t)_+^{m-1}, \quad (4)$$

for  $k = -m + 1, \dots, 0, 1, 2, \dots$

Numerically, we would apply the recursive formula to implement  $N_{m,k}$  (see, for example, [15, page 143 (6.6.12a)]). It is also shown in [24] that  $\{N_{\underline{s},m,k}\}$ ,  $k = -m + 1, \dots, 0, 1, 2, \dots$  is a stable locally supported basis of the space  $S_{\underline{s},m}$  in that every function  $f \in S_{\underline{s},m} \cap L^\infty[a, \infty)$  has a unique spline series representation with coefficient sequence  $\{c_k\} \in \ell^\infty$ . Moreover, if  $\delta_m := \inf_k(t_{m+k} - t_k) > 0$ , then

$$c(m, \delta_m) \|\{c_k\}\|_{\ell^\infty} \leq \left\| \sum_{k=-m+1}^{\infty} c_k N_{\underline{s},m,k}(t) \right\|_{L^\infty[a, \infty)} \leq \|\{c_k\}\|_{\ell^\infty}, \quad (5)$$

where  $c(m, \delta_m) > 0$  is a constant depending only on  $m$  and the “minimum knot spacing”  $\delta_m$ .

However, if we choose the sequence  $\underline{t}$  of time positions in (2) as the knot sequences  $\underline{s}$  in (3) by attaching  $t_{-m+1} = \dots = t_0 = a$  to  $\underline{t}$  as in (3), there are possibly two complications. Firstly, for the spline space  $S_{\underline{t},m}$  of odd order  $m \geq 5$ , computation of the spline interpolant

$$f(t) = \sum_k c_k N_{\underline{t},m,k}(t)$$

of the “clean” non-uniform samples  $\hat{x}_k = g(t_k)$ , where  $k = 0, 1, \dots, n$ , could be unstable, since the coefficient matrix

$$A_{n+1} := [N_{\underline{t},m,k}(t_j)]_{j=0,\dots,n; k=-m+1,\dots,n-m+1}$$

may not be diagonal dominant. Secondly, for even order  $m \geq 4$ , although the matrix  $A_{n+1}$  is usually diagonal dominant, the computational cost to obtain the coefficients  $c_{-m+1}, c_{-m+2}, \dots, c_{n-m+1}$  is high for non-uniform  $\{t_j\}$  and large values of  $n$ .

On the other hand, if the spline representer  $f(t)$  is not required to interpolate the target data function  $g(t)$  at  $t = t_j$ ,  $j = 0, 1, \dots, n$  (that is, if  $f(t_j) \neq g(t_j)$  is allowed), then the “quasi-interpolation” scheme introduced by de Boor and Fix [25] meets the requirement of “polynomial preservation”, in that for  $g(t) = p(t) \in \Pi_{m-1}$  in a desired neighborhoods of some time sample  $t^*$ , then  $f(t) = p(t)$  in this neighborhood. Polynomial preservation is important to facilitate the continuous wavelet transform (CWT), and hence the synchrosqueezing transform (SST) for oscillatory components separation and analysis, such as analysis of the respiratory signal from the ECG signal as will be discussed in Section 4 in this paper. In particular, if the analysis wavelet is also a compactly supported spline of order  $m$ , then the  $m$ -th order vanishing moment annihilates the  $m$ -th order Taylor polynomial expansion of  $g(t)$ , that facilitates in removing the trend of the signal  $g(t)$ . However, the quasi-interpolation scheme in [25, 24] requires derivative data values of  $g(t)$  that are not available in our study. In [42] and [50], derivatives of  $g(t)$  are replaced by divided differences of  $\{g(t_i)\}$ , for our need. Other quasi-interpolation scheme including a comprehensive study in [26] have also been derived. But to the best of our knowledge, only the basis functions introduced in our work [10] have the real-time approximation feature. As mentioned above, quasi-interpolation introduced the error  $g(t_i) - f(t_i)$  for  $i = 0, 1, \dots, n$ . In [18], the notion of local interpolation is introduced to correct the errors, leading to the “blending operator”, to be discussed in the following sub-section.

### 2.1. Blending operator

Let us first study the real-time quasi-interpolant introduced in [10]. We start from preparing some notations. For each  $k = 0, 1, \dots$ , consider the Vandermonde determinant

$$D(t_k, \dots, t_{k+m+1}) := \det \begin{bmatrix} 1 & 1 & \dots & 1 \\ t_k & t_{k+1} & \dots & t_{k+m-1} \\ \vdots & \vdots & \ddots & \vdots \\ t_k^{m-1} & t_{k+1}^{m-1} & \dots & t_{k+m-1}^{m-1} \end{bmatrix}$$

and the determinant  $D(t_k, \dots, t_{k+j-1}, \xi_j, t_{k+j+1}, \dots, t_{k+m-1})$  obtained by replacing the  $(j+1)$ -st column in the definition of  $D(t_k, \dots, t_{k+m+1})$  by the column vector

$$\xi_j := [\xi^0(j, m), \dots, \xi^{m-1}(j, m)]^T,$$

where  $\xi^0(j, m) = 1$  and

$$\xi^i(j, m) = \frac{\sigma^i(t_{j+1}, \dots, t_{j+m-1})}{\binom{m-1}{i}}$$

for  $i = 1, \dots, m-1$  with  $\sigma^i(r_1, \dots, r_{m-1})$  being the classical symmetric functions defined by  $\sigma^0(r_1, \dots, r_{m-1}) = 1$  and for  $i = 1, \dots, m-1$ ,

$$\sigma^i(r_1, \dots, r_{m-1}) = \sum_{1 \leq l_1 < \dots < l_i \leq m-1} r_{l_1} \dots r_{l_i}.$$

**Definition** [Quasi-interpolation operator] Using the determinants introduced above, we apply the spline coefficients

$$a_{k,l} := \frac{D(t_k, \dots, t_{k+j-1}, \xi_j, t_{k+j+1}, \dots, t_{k+m-1})}{D(t_k, \dots, t_{k+m-1})}$$

to formulate the compactly supported spline function

$$M_{\underline{t},m,k}(t) := \sum_{l=m-1}^{2m-2} a_{k,l-m+1} N_{\underline{t},m,k+l-m+1}(t)$$

with  $\text{supp} M_{\underline{t},m,k} = [t_{k-m+1}, t_{k+m}]$ . These basis functions provide a real-time implementation of the *quasi-interpolation operator*

$$(\mathbf{Q}_m g)(t) = \sum_k g(t_k) M_{\underline{t},m,k}(t). \quad (6)$$

We summarize some of the properties of the quasi-interpolation operator  $\mathbf{Q}_m$  in (6).

**Lemma 2.1.** *For any  $m \geq 1$ ,  $\mathbf{Q}_m$  possesses the polynomial preservation property*

$$(\mathbf{Q}_m p)(t) = p(t)$$

for all  $t \geq a$  and for all  $p \in \Pi_{m-1}$ , provided that the summation in (6) is taken for all non-negative integers  $k = 0, 1, \dots$ . Furthermore, in view of the support of  $M_{\underline{t},m,k}$ , it follows that

$$\sum_{k=v-m+1}^{v+m-2} p(t_k) M_{\underline{t},m,k}(t) = p(t), \quad t \in [t_u, t_v]$$

for all  $p \in \Pi_{m-1}$ .

This local polynomial preservation property allows the CWT, and hence the SST, to annihilate the  $(m-1)$ -th degree Taylor polynomial approximation of the signal at  $t_j$ , where  $u < j < v$ .

We now turn to the discussion of the local interpolation operator, to be denoted by  $\mathbf{R}_m$ , which satisfies the interpolation property. To define  $\mathbf{R}_m$ , we may insert knots to  $\underline{t}$  by considering a new knot sequence  $\underline{s} \supset \underline{t}$  as in (3). More precisely, we consider even and odd orders separately, as follows:

(i) For even  $m \geq 4$ , we set

$$s_{mk/2} = t_k, \quad k = 0, 1, 2, \dots$$

That is, we insert  $(m/2-1)$  knots in between two consecutive knots in  $\underline{t}$ . For convenience, we may choose the new knots equally spaced in-between every pair of two consecutive knots.

**Definition** [Completely local spline basis function] Fix even  $m \geq 4$ . Let  $N_{\underline{s},m,j}$  be the  $m$ -th order B-spline with knot sequence  $\underline{s}$ . Then the *completely local spline basis function* can be defined by

$$L_{\underline{s},m,j}(t) := \frac{N_{\underline{s},m,m(j-1)/2}(t)}{N_{\underline{s},m,m(j-1)/2}(t_j)}.$$

Since  $t_j = s_{mj/2}$  is the “centered” knot and

$$\text{supp} N_{\underline{s},m,m(j-1)/2} = [s_{m(j-1)/2}, s_{m(j+1)/2}] = [t_{j-1}, t_{j+1}],$$

we have

$$\begin{cases} L_{\underline{s},m,j}(t_j) = 1 \\ \text{supp} L_{\underline{s},m,j} = [t_{j-1}, t_{j+1}]. \end{cases} \quad (7)$$

In view of (7), it is clear that  $L_{\underline{s},m,j}(t_k) = \delta_{j-k}$ , where the Kronecker delta notation is used.

**(ii)** For odd  $m \geq 3$ , we may insert  $\frac{m+1}{2} - 1$  equally spaced new knots in between  $[t_{2k}, t_{2k+1}]$ , and  $\frac{m+1}{2} - 2$  equally spaced new knots in between  $[t_{2k+1}, t_{2k+2}]$ , for  $k = 0, 1, \dots$ . Then, by considering even and odd indices ( $j = 2k$  and  $j = 2k + 1$ ) separately, a similar construction as the even order setting yields completely local spline basis functions  $L_{\underline{s},m,j}(t)$  that has the property (7) as the even order  $m$ .

The above preparation provides a real-time implementation of the *local interpolation operator*, which satisfies the interpolation property due to (7).

**Definition** [Local interpolation operator] Fix  $m \geq 3$ . For a given function  $g \in C(\mathbb{R})$ , the local interpolation operator  $R_m$  is defined by

$$(R_m g)(t) := \sum_k g(t_k) L_{\underline{s},m,k}(t). \quad (8)$$

We are now ready to apply (6) to obtain the blending operator, denoted by  $R_m \oplus Q_m$ .

**Definition** [Blending operator] Fix  $m \geq 3$  and  $g \in C(\mathbb{R})$ . The blending operator is defined as  $P_m := R_m \oplus Q_m$ , where

$$R_m \oplus Q_m := Q_m + R_m(I - Q_m) = Q_m + R_m - R_m Q_m,$$

and  $I$  is the identity operator. In particular, we have

$$(P_m g)(t) := \sum_k g(t_k) M_{\underline{t},m,k}(t) + \sum_k \left[ g(t_k) - \sum_j g(t_j) M_{\underline{t},m,j}(t_k) \right] L_{\underline{s},m,k}(t). \quad (9)$$



We remark that in the definition of  $P_m$ , the two operators  $R_m$  and  $Q_m$  are not commutative. Let us summarize the two key properties of the blending operator in the following theorem.

**Theorem 2.2.** *The blending operator  $P_m$  possesses both the polynomial preservation property of  $Q_m$  and the interpolatory property of  $R_m$ .*

*Proof.* Indeed, for any  $p \in \Pi_{m-1}$ , since  $Q_m p = p$ , we have

$$(R_m \oplus Q_m)p = Q_m p + R_m p - R_m(Q_m p) = p + R_m p - R_m p = p.$$

Also,  $f := (R_m \oplus Q_m)g \in S_{\underline{t},m}$  interpolates the discrete data  $\{g(t_j)\}$  in that

$$f(t_j) = (Q_m g)(t_j) + (R_m g)(t_j) - R_m(Q_m g)(t_j) = (Q_m g)(t_j) + g(t_j) - (Q_m g)(t_j) = g(t_j).$$

Observe that the definition of  $(P_m g)(t)$  in (9) depends only on the discrete samples  $g(t_k)$ ,  $k = 0, 1, \dots$  of  $g(t)$  and that

$$f(t) := (P_m g)(t) \in S_{\underline{s},m}[0, \infty)$$

since  $\underline{t} \subset \underline{s}$ . As already discussed earlier, we have, firstly, the interpolatory property

$$f(t_j) = g(t_j), \quad j = 0, 1, 2, \dots$$

and secondly, the polynomial preservation property

$$(P_m p)(t) = p(t), \quad t \geq a, \quad (10)$$

for all  $p \in \Pi_{m-1}$ . In fact, if only  $t \in [t_u, t_v]$  in (10) is desired, then the summation over  $k$  in (9) is taken from  $k = u - m + 1$  to  $k = v + m - 2$ , due to the finite support property of  $M_{\underline{t},m,k}(t)$  and  $L_{\underline{s},m,k}(t)$ .  $\square$

To study how well the blending operator approximates functions that are not polynomials, we consider only knot sequences  $\underline{t}$  that satisfies the condition

$$\frac{\sup_{0 \leq i \leq k} |t_{k+i+m} - t_{k+i}|}{\inf_{0 \leq i \leq k} |t_{k+i+1} - t_{k+i}|} \leq C_{\underline{t},k} \quad (11)$$

for some positive constant  $C_{\underline{t},k}$ , where  $m > 0$  is the order of the B-splines  $N_{\underline{t},m,k}$  and  $k \geq 0$  is the knot index. Under this condition, it has been shown in [10, Lemma 2,1] that the coefficients  $a_{k,j}$  in the definition of the quasi-interpolation operator (6) (with  $g(t_j)$  replaced by  $a_{k,j}$ ) satisfy

$$|a_{k,j}| \leq \frac{1}{(m-1)!} C_{\underline{t},k}.$$

Since the B-splines  $N_{\underline{t},m,j}$  are non-negative and constitute a partition of unity, we have

$$|M_{\underline{t},m,k}(t)| \leq \frac{1}{(m-1)!} C_{\underline{t},k}^{m-1} \sum_{l=m-1}^{2m-2} N_{\underline{t},m,k+l-m+1}(t) \leq \frac{1}{(m-1)!} C_{\underline{t},k}^{m-1} \quad (12)$$

for all  $t \geq a$ . With this preparation, we are ready to formulate our result on error estimation.

**Theorem 2.3.** Fix  $m > 1$ . Let  $s$  be an arbitrary integer with  $0 \leq s \leq m - 1$ , and  $M_{\underline{t},m,k}(t)$  be the quasi-interpolation spline basis functions defined in (6) in terms of the  $m$ -th order B-splines on a knot sequence  $\underline{t}$  that satisfies (11). Then for a chosen  $k \geq 0$ , there exists a constant  $C_{m,s,k} > 0$  such that for every  $g \in C^{s+1}[a, \infty)$ , the error of spline interpolation at the knots  $t_j$  by the blending operator is given by

$$\|g - P_m g\|_{L^\infty[t_k, t_{k+1}]} \leq C_{m,s,k} \|g^{(s+1)}\|_{L^\infty[t_k, t_{k+1}]} \delta_{\underline{t},k}^{s+1},$$

where

$$C_{m,s,k} := \frac{1}{s!} \left( 2 + \frac{m(m-2)}{(m-1)!} C_{\underline{t},k}^{m-1} \right) \quad (13)$$

and

$$\delta_{\underline{t},k} := \max_{k+1 \leq i \leq k+m} (t_{i+1} - t_i).$$

*Proof.* For any fixed  $k \geq 0$ , consider the Taylor polynomial expansion of  $g$  at  $t_k$

$$g(t) = \sum_{j=0}^s \frac{g^{(j)}(t_k)}{j!} (t - t_k)^j + \int_{t_k}^{t_{k+1}} \frac{g^{(s+1)}(\tau)}{s!} h(t, \tau) d\tau,$$

where  $t \in [t_k, t_{k+1}]$  and  $h(t, \tau) := (t - \tau)_+^s$ . Since the blending operator  $P_m$  preserves all polynomials of degree  $\leq m - 1$  and  $s \leq m$  is arbitrary, we have

$$g(t) - P_m g(t) = \int_{t_k}^{t_{k+1}} \frac{g^{(s+1)}(\tau)}{s!} [h(t, \tau) - (P_m h(\cdot, \tau))(t)] d\tau.$$

In the following, we suppress the variable of integration  $\tau$  by writing  $h(t) = h(t, \tau)$ , so that

$$\begin{aligned} h(t) - (P_m h)(t) &= h(t) - \sum_{j=k-m+1}^{k+m-2} h(t_j) M_{\underline{t},m,j}(t) - \sum_{j=k-1}^{k+1} \left[ h(t_j) - \sum_{l=j-m+1}^{j+m-2} h(t_l) M_{\underline{t},m,l}(t_j) \right] L_{\underline{s},m,j}(t). \end{aligned}$$

Now, observe that since  $h(t_j) = (t_j - \tau)_+^s$  and  $\tau \in [t_k, t_{k+1}]$ , we have  $h(t_j) = 0$  for all  $j \leq k$ . Hence,

$$\begin{aligned} h(t) - (P_m h)(t) &= h(t) - \sum_{j=k+1}^{k+m-2} h(t_j) M_{\underline{t},m,j}(t) - h(t_{k+1}) L_{\underline{s},m,k+1}(t) + \sum_{l=k+1}^{k+m-1} h(t_l) M_{\underline{t},m,l}(t_{k+1}) L_{\underline{s},m,k+1}(t). \end{aligned}$$

From (12) and the fact that  $|L_{\underline{s},m,j}(t)| \leq 1$  for all  $t \in [t_k, t_{k+1}]$ , we have

$$\begin{aligned} |h(t) - (P_m h)(t)| &\leq 2\delta_{\underline{t},k}^s + \frac{1}{(m-1)!} C_{\underline{t},k}^{m-1} (\delta_{\underline{t},k}^s + 2\delta_{\underline{t},k}^s + \dots + (m-2)\delta_{\underline{t},k}^s) \\ &\quad + \frac{1}{(m-1)!} C_{\underline{t},k}^{m-1} (\delta_{\underline{t},k}^s + 2\delta_{\underline{t},k}^s + \dots + (m-1)\delta_{\underline{t},k}^s) \\ &= \left( 2 + \frac{m(m-2)}{(m-1)!} C_{\underline{t},k}^{m-1} \right) \delta_{\underline{t},k}^s. \end{aligned}$$

Therefore, it follows that

$$\begin{aligned}\|g - \mathbf{P}_m g\|_{L^\infty[t_k, t_{k+1}]} &\leq \int_{t_k}^{t_{k+1}} \frac{g^{(s+1)}(\tau)}{s!} \left(2 + \frac{m(m-2)}{(m-1)!} C_{\underline{t}, k}^{m-1}\right) \delta_{\underline{t}, k}^s d\tau \\ &\leq C_{m, s, k} \|g^{(s)}\|_{L^\infty[t_k, t_{k+1}]} \delta_{\underline{t}, k}^{s+1}.\end{aligned}$$

□

In conclusion, the blending operator, as a local spline interpolation operator, achieves the optimal interpolation error rate compared with the traditional spline interpolation operator. In addition, the error depends only on the local data profile, which allows the real-time implementation with the optimal error rate.

## 2.2. Real-time implementation

We conclude our discussion by giving a brief description of a real-time computational scheme to compute  $f(t)$  for the in-coming data samples  $g(t_0), g(t_1), \dots$ . For convenience, we only consider even order  $m$ . The formulation for the odd order is similar but slightly more complicated.

For uniform sampling  $\{g(k\tau)\}$ ,  $k = 0, 1, 2, \dots$ , where  $\tau > 0$ , a very simple algorithm is derived in [14, p. 114-117].

**Remark** For larger values of the spline order  $m$ , observe that the support of the completely local spline basis function  $L_{\underline{s}, m, j}(x)$  could be too small, perhaps resulting in some undesirable “bumpy” shape of the interpolant  $\mathbf{P}_m g$  of the data function  $g$ . Therefore, at the expense of higher computational cost and slower on-line performance, local interpolants with larger support, such as approximate modification of those constructed in [17] could be constructed to replace the simple interpolants  $L_{\underline{s}, m, j}(x)$ .

## 3. Vanishing-moment (VM) wavelets with minimum support

In this section, we introduce the notion of vanishing-moment (VM) wavelets on bounded intervals. In Subsection 3.1, we first disclose the method of construction of those with minimum support (and maximum order of vanishing moments) in terms of the B-splines  $N_{\underline{s}, m, k}(t)$  of arbitrary order  $m \geq 2$  on the knot sequence  $\underline{s}$  in (3), with the following modification:

$$\underline{s} : s_{-m+1} = \dots = s_0 = 0 < s_1 < \dots < N = s_N = \dots = s_{N+m-1} \quad (14)$$

where  $s_j = j$ , for  $j = 0, \dots, N$ . More precisely, we will first introduce the wavelets  $\psi_{\underline{s}, m, j}(x)$ ,  $j = -m+1, \dots, 2N-m-1$ , that, up to any constant multiple, are uniquely determined by the following governing conditions:

$$\psi_{\underline{s}, m, n, j}(x) := \sum_{k=j}^{n+j} q_{n, j, k} N_{\frac{1}{2}\underline{s}, m, k}(x) \quad (15)$$

---

**Algorithm 1** Real-time implementation of the blending operator

---

First, pre-compute the B-spline values  $N_{\underline{s},m,l}(t_j) =: n_{l,j}$ . Then for each  $k$ , since  $M_{\underline{t},m,k} \in S_{\underline{s},m}$ , there exists a finite sequence  $\{b_{k,l}\}$  in the formulation of

$$M_{\underline{t},m,k}(t) = \sum_l b_{k,l} N_{\underline{s},m,l}(t).$$

Also pre-compute  $d_{k,j} = M_{\underline{t},m,k}(t_j)$ .

Now, while the data sequence  $\{g(t_k)\}$  is acquired, compute

$$\tilde{g}_l = \sum_k b_{k,l} g(t_k)$$

and simultaneously compute

$$g_l^* = \frac{g(t_l) - \sum_j d_{l,j} g(t_j)}{n_{m(l-1)/2, l}},$$

and then up-sample  $\{g_l^*\}$  by  $m(l-1)/2$ ; that is, set

$$g_{m(k-1)/2}^\# = g_k^*, \quad \text{and } g_l^\# = 0 \text{ otherwise.}$$

Then, we have an on-line computational scheme for the quasi-interpolation spline interpolation:

$$f(t) = \sum_{l=-m+1}^n (\tilde{g}_l - g_l^\#) N_{\underline{s},m,l}(t)$$

for increasing number of samples from  $g(t_n)$  to  $g(t_{n+1}), \dots$ ; and this can be implemented for real-time D/A conversion.

---

and

$$\int_0^N x^l \psi_{\tilde{\underline{s}},m;n,j}(x) dx = 0, \quad \text{for } l = 0, 1, \dots, n-1, \quad (16)$$

where the knot sequence  $\frac{1}{2}\underline{s}$  is defined by

$$\tilde{\underline{s}} := \frac{1}{2}\underline{s} : \tilde{s}_{-m+1} = \dots = \tilde{s}_0 = 0 < \tilde{s}_1 < \dots < \tilde{s}_{2N-1} < \tilde{s}_{2N} = 2N = \dots = \tilde{s}_{2N+m-1}, \quad (17)$$

with  $\tilde{s}_j = \frac{1}{2}j$ , for  $j = 0, \dots, 2N$ , by “halving” the knot sequence  $\underline{s}$ , but with extension from  $j = N$  to  $j = 2N$  to fill the given bounded interval  $[0, N]$ .

Observe that in view of the  $n$ -th order moment condition in (16), the  $n+1$  coefficients  $q_{n;j,k}$ ,  $k = j, \dots, j+n$ , in (15), we are assured that the non-trivial VM wavelets  $\psi_{\tilde{\underline{s}},m;n,j}(x)$  are unique, up to a non-zero constant multiple. Indeed, since  $N_{\tilde{\underline{s}},m,k}(x)$  reproduces polynomials [14] and there are  $n$  equations, when  $n \leq m$ , the solution to the  $n+1$  coefficients are free up to  $\mathbb{R}$ . In addition, the VM wavelets  $\psi_{\tilde{\underline{s}},m;n,j}(x)$  have minimum support:

$$\text{supp } \psi_{\tilde{\underline{s}},m;n,j} = [\tilde{s}_j, \tilde{s}_{j+m+n}], \quad j = -m+1, \dots, 2N-m-1,$$

so that the support of the “interior wavelets”, with  $j = 0, \dots, 2N - 2m$ , are given by

$$\text{supp } \psi_{\underline{s}, m, j} = \left[ \frac{j}{2}, \frac{j + m + n}{2} \right].$$

Since (almost) all wavelets in the literature embrace the multi-resolution analysis (MRA) structure and are generated by the operations of translation and dilation, we will first construct our VM (spline) wavelets on uniform knots in Subsection 3.1. In addition, since (almost) all spline wavelets introduced in the wavelet literature restrict the order (i.e. number)  $n$  of vanishing moments of the wavelets not to exceed the order  $m$  of the spline function (i.e.  $n \leq m$ ), we only consider  $n = m$  in Subsection 3.1. After all, VM wavelets for the MRA setting could have their place in further theoretical development and applications in the future. However, for our application of wavelets to the synchrosqueezed continuous wavelet transform in this paper, the MRA structure is not required. In Subsection 3.2, we will extend the VM spline wavelets to spline functions on arbitrary knot sequences and with arbitrary order  $n$  of vanishing moments, and allow  $n > m$ . For example, for  $m = 1$ , it follows from our general result in Theorem 3.1 that a compactly supported piecewise constant wavelet  $\psi_{\underline{s}, 1, n}(x)$ , with finitely many jump discontinuities on an arbitrarily chosen knot sequence  $\underline{s}$ , can be constructed to have any desirable number  $n$  of vanishing moments.

### 3.1. VM wavelets with the same number of vanishing moments as the spline order

This subsection is focused on MRA vanishing moment wavelets on a bounded interval. The general VM wavelets will be discussed in Subsection 3.2. Let us first consider the special case where the order  $n$  of vanishing moments agrees with that of the VM spline wavelets. This is the initial investigation that leads to the general theory to be discussed in Subsection 3.2. In addition, for our application to the SST, to be studied in Section 4, these are the preferred wavelets on a bounded interval. For  $n = m$ , we simplify the notation by setting

$$\psi_{\underline{s}, m, j}(x) := \psi_{\underline{s}, m; m, j}(x); \quad q_{j, k} := q_{n; j, k}.$$

Observe that the length of the support of the interior wavelets  $\psi_{\underline{s}, m, j}$  (i.e. for  $j = 0, \dots, 2N - 2m$ ) is the same as the length of the support of the interior B-splines  $N_{\underline{s}, m, k}(x) = N_{\underline{s}, m, 0}(x - j)$ ,  $j = 0, \dots, 2N - m - 1$ , that generate the interior VM wavelets. There are  $2(m - 1)$  boundary VM wavelets, with  $\psi_{\underline{s}, m, j}$ ,  $j = -m + 1, \dots, -1$  for the end-point  $s_0 = 0$ , and  $\psi_{\underline{s}, m, j}$ ,  $j = 2N - 2m + 1, \dots, 2N - m - 1$  for the other end-point  $s_N = \tilde{s}_{2N} = 0$ . It follows from the formulation (15) for these boundary wavelets that their supports are

$$\text{supp } \psi_{\underline{s}, m, j} = \left[ 0, \frac{2m + j}{2} \right], \quad j = -m + 1, \dots, -1,$$

and

$$\text{supp } \psi_{\underline{s}, m, j} = \left[ \frac{j}{2}, N \right], \quad j = 2N - 2m + 1, \dots, 2N - m - 1.$$

For each  $j = -m + 1, \dots, 2N - m - 1$ , to compute the sequence

$$\underline{q}_j := \{q_{j, j}, \dots, q_{j, m+j}\},$$

we may simply apply (16) to obtain the homogeneous system of  $m$  linear equations with  $m + 1$  unknowns in  $\underline{q}_j$ :

$$\sum_{k=j}^{m+j} \left( \int_0^N x^l N_{\tilde{\mathbf{s}},m,k}(x) dx \right) q_{j,k} = 0;$$

or equivalently,

$$\sum_{k=j}^{m+j} \left( \int_0^N x^l N_{\underline{\mathbf{s}},m,k}(x) dx \right) q_{j,k} = 0; \quad (18)$$

which clearly has the same solution (or null) space. We next introduce the family  $\{\underline{c}^l\}$ ,  $l = 0, \dots, m-1$ , of sequences

$$\underline{c}^l := \{c_{-m+1}^l, \dots, c_{N-1}^l\},$$

defined by

$$c_j^l := (-1)^l \frac{l!}{(m-1)!} \left[ \frac{d^{m-l-1}}{dt^{m-l-1}} \prod_{i=1}^{m-1} (t - s_{i+j}) \right] \Big|_{t=0}.$$

Then it follows from Marsden's identity [43] that

$$x^l = \sum_{u=-m+1}^{N-1} c_u^l N_{\underline{\mathbf{s}},m,u}(x), \quad x \in [0, N],$$

for  $l = 0, \dots, m-1$  (see [13, p.6 Eq(1.12)]). Therefore, the linear system (18) can be reformulated as

$$\sum_{k=j}^{m+j} d_k^l q_{j,k} = 0, \quad l = 0, \dots, m-1, \quad (19)$$

where

$$d_k^l := \sum_{u=-m+1}^{N-1} c_u^l \langle N_{\underline{\mathbf{s}},m,k}, N_{\underline{\mathbf{s}},m,u} \rangle. \quad (20)$$

Since it can be shown that each of the  $(m+1) \times m$  matrices

$$A_j = [d_k^l]_{j \leq k \leq m+j, 0 \leq l \leq m-1},$$

where  $j = -m+1, \dots, 2N-m-1$ , has full rank  $m$ , the null space of (19) has dimension 1; that is, the VM wavelets  $\psi_{\underline{\mathbf{s}},m,j}(x)$ ,  $j = -m+1, \dots, 2N-m-1$ , in (15) are unique, up to a (non-zero) constant multiple.

The remaining computational scheme to be discussed is the computation of the inner products in (20). For convenience, we drop the subscript  $\underline{\mathbf{s}}$  in the B-spline; that is,  $N_{m,k} = N_{\underline{\mathbf{s}},m,k}$ . For interior B-spline,  $N_{m,k}$  and  $N_{m,u}$ , that is,  $0 \leq k, u \leq N-m$ , these B-splines are simply integer translations of the B-spline  $N_{m,0}(x)$ , so that

$$\begin{aligned} \langle N_{m,k}, N_{m,u} \rangle &= \int_0^N N_{m,0}(x-k) N_{m,0}(x-u) dx \\ &= \int_0^N N_{m,0}(x) N_{m,0}(x-(u-k)) dx \\ &= \int_0^N N_{m,0}(m-x) N_{m,0}(x-(u-k)) dx \end{aligned}$$

since  $N_{m,0}(m-x) = N_{m,0}(x)$  by the symmetric property of  $N_{m,0}(x)$ . Hence, by the convolution property of cardinal B-spline with integer knots, we have

$$\langle N_{m,k}, N_{m,u} \rangle = \int_{-\infty}^{\infty} N_{m,0}((m-u+k)-x)N_{m,0}(x)dx = N_{2m,0}(m-u+k). \quad (21)$$

There are various methods for evaluating the cardinal B-spline  $N_{2m,0}(x)$  at integers  $x = v := m - u + k$ . The most popular one is “weighted differencing” of two lower order B-spline, namely,

$$\begin{aligned} N_{2m,0}(v) &= \frac{1}{2m-1} [vN_{2m-1,0}(v) + (2m-v)N_{2m-1,0}(v-1)] \\ N_{2m-1,0}(w) &= \frac{1}{2m-2} [wN_{2m-2,0}(w) + (2m-1-w)N_{2m-2,0}(w-1)] \end{aligned}$$

for  $v, w \in \mathbb{Z}$ , and so forth, till we arrive at

$$N_2(v) = \delta(v), \quad v \in \mathbb{Z};$$

(see, for example, [14, p.86, Eq(4.2.4)]). Another, perhaps more efficient, algorithm is the algorithm by applying the generalized Pascal triangle as described in [14, p.190-p.191]. For example, for  $m = 4$  and  $m = 5$ , the non-zero values of  $N_{2m,0}(v)$ ,  $v \in \mathbb{Z}$ , are listed in row #7 and row #9, respectively, in [14, p.191, Table 6.3.1], namely

1.  $7!N_8(v) = 1, 120, 1191, 2416, 1191, 120, 1$ , for  $v = 1, \dots, 7$  respectively.
2.  $9!N_{10}(v) = 1, 502, 14608, 88234, 156190, 88234, 14608, 502, 1$ , for  $v = 1, \dots, 9$  respectively.

On the other hand, if one of the two B-splines  $N_{m,k}$  and  $N_{m,u}$  in the inner product  $\langle N_{m,k}, N_{m,u} \rangle$  is not an interior B-spline, then the convolution property is not valid. In this case, the algorithm developed in [19] (see also [13, p.12]) for computing each polynomial piece of the B-spline  $N_{m,k}(u)$  on the interval  $[j, j+1]$ ,  $j = k, \dots, k+m-1$ , in terms of Bernstein polynomials can be applied. We remark that an efficient scheme for computing the inner product on  $[0, 1]$  of two Bernstein polynomials of the same degree is available in the literature.

**Example** [VM wavelets  $\psi_{\underline{g},4,j}(x)$  in (15) with  $m = 4$ ]

1. For the interior VM wavelets; that is, for  $j = 0, \dots, 2N-8$ , the coefficient sequence  $\underline{q}_j = \{q_{j,j}, \dots, q_{j,j+4}\}$  is given by any non-zero constant multiple of  $\{1, -4, 6, -4, 1\}$ .
2. For the  $m-1 = 4-1 = 3$  boundary VM wavelets corresponding to the end-point  $x = 0$ , we only list  $\psi_{\underline{g},4,-2}(x)$  and  $\psi_{\underline{g},4,-1}(x)$  since we do not need  $\psi_{\underline{g},4,-3}(x)$  which does not vanish at  $x = 0$ .

(a)  $\underline{q}_{-1} = \{q_{-1,-1}, q_{-1,0}, q_{-1,1}, q_{-1,2}, q_{-1,3}\}$  is any non-zero constant multiple of

$$\left\{ \frac{7}{3}, \frac{-319}{60}, \frac{101}{15}, \frac{-25}{6}, 1 \right\}.$$

(b)  $\underline{q}_{-2} = \{q_{-2,-2}, q_{-2,-1}, q_{-2,0}, q_{-2,1}, q_{-2,2}\}$  is any non-zero constant multiple of

$$\left\{ 1, \frac{-116}{25}, \frac{919}{100}, \frac{-57}{5}, 6 \right\}.$$

3. For the boundary VM wavelet corresponding to the end-point  $x = N$ , we apply the symmetry property to write

$$\begin{aligned}\psi_{\tilde{\mathbf{s}},4,2N-9}(x) &= \psi_{\tilde{\mathbf{s}},4,-1}(N-x), \\ \psi_{\tilde{\mathbf{s}},4,2N-8}(x) &= \psi_{\tilde{\mathbf{s}},4,-2}(N-x), \\ \psi_{\tilde{\mathbf{s}},4,2N-7}(x) &= \psi_{\tilde{\mathbf{s}},4,-3}(N-x);\end{aligned}$$

but, again, since  $\psi_{\tilde{\mathbf{s}},4,2N-7}(x)(N) = \psi_{\tilde{\mathbf{s}},4,-3}(0)$  does not vanish, we only need  $\psi_{\tilde{\mathbf{s}},4,2N-9}(x)$  and  $\psi_{\tilde{\mathbf{s}},4,2N-8}(x)$  in our application.

Motivated by the above example of interior VM spline-wavelets  $\psi_{\tilde{\mathbf{s}},4,j}(x)$  with  $m = 4$ , let us consider the general setting

$$\psi_m(x) := \sum_{k=0}^m (-1)^k \binom{m}{k} N_m(2x - k), \quad (22)$$

for an arbitrary integer  $m \geq 1$ , where

$$N_m(x) := N_{m,0}(x) = \int_0^1 N_{m-1}(x-t) dt, \quad (23)$$

defined recursively, with  $N_1(x) = \chi_{[0,1)}(x)$ . Observe that  $\psi_1(x) = N_1(2x) - N_1(2x-1)$  is the Haar wavelet.

**Theorem 3.1.** *For any integer  $m \geq 2$  and  $j = 0, \dots, 2N-2m$ , the VM wavelets defined in (15) and (16) are (any non-zero constant multiple of)*

$$\psi_{m,j}(x) = \psi_m(x - j/2), \quad 0 \leq j \leq 2(N-m),$$

(called the interior VM wavelets), where  $\psi_m(x)$  is defined in (22). Furthermore, the Fourier transform of  $\psi_m(x)$  is given by

$$\hat{\psi}_m(\omega) = \frac{1}{2} (i\omega/2)^m \hat{N}_{2m}(\omega/2); \quad (24)$$

or equivalently,

$$\psi_m(x) = N_{2m}^{(m)}(2x). \quad (25)$$

*Proof.* The formula (23) for the so-called *cardinal B-spline*  $N_{m,0}$  is well-known (and in fact is commonly used to define  $N_m(x)$ ). Now, since the family  $\{q_{j,k}\}$  of finite sequences in the definition of the VM wavelets in (15) and (16) is unique (up to a non-zero constant multiples), and since

$$\psi_{\tilde{\mathbf{s}},m,j}(x) = \psi_{\tilde{\mathbf{s}},m,0}(x - j/2)$$

for the interior VM wavelets (that is, for  $j = 0, \dots, 2(N-m)$ ), it is sufficient to prove that  $\psi_m(x)$ , as defined in (22), meets the vanishing moment requirement (16); that is,

$$\int_{-\infty}^{\infty} x^l \psi_m(x) dx = 0, \quad l = 0, \dots, m-1. \quad (26)$$



Observe that from the definition (22), it is clear that  $\psi_m(x)$  has compact support with  $\text{supp } \psi_m = [0, m]$ , so that its Fourier transform  $\hat{\psi}_m(\omega)$ , along with the derivatives  $D^l \hat{\psi}_m(\omega)$ ,  $l = 1, 2, \dots$ , are well-defined. Hence, since (26) is equivalent to

$$(D^l \hat{\psi}_m)(0) = 0, \quad l = 0, \dots, m-1, \quad (27)$$

it is sufficient to verify (27). To compute the Fourier transform of  $\psi_m(x)$ , in (22), we observe that  $\hat{N}_m(\omega) = (\hat{N}_1(\omega))^m$ , so that

$$\begin{aligned} \hat{\psi}_m(\omega) &= \sum_{k=0}^m (-1)^k \binom{m}{k} \frac{1}{2} e^{-ik\omega/2} \hat{N}_m(\omega/2) \\ &= \frac{1}{2} (1 - e^{-i\omega/2})^m (\hat{N}_1(\omega/2))^m \\ &= \frac{1}{2} (1 - e^{-i\omega/2})^m \left( \frac{1 - e^{-i\omega/2}}{i\omega/2} \right)^m \\ &= \frac{1}{2} (i\omega/2)^m \left( \frac{1 - e^{-i\omega/2}}{i\omega/2} \right)^{2m} \\ &= \frac{1}{2} (i\omega/2)^m \hat{N}_{2m}(\omega/2). \end{aligned}$$

Therefore, if  $\omega$  is extended to a complex variable, then  $\hat{\psi}_m(\omega)$  is an entire function with  $m$ -fold zero at  $\omega = 0$ . This completes the proof of (26), and hence of (27). The above computation also yields (24). The proof of (25) is simply integration of the Fourier transform of  $N_{2m}^{(m)}(2x)$  by parts  $m$  times, followed by application of (24).  $\square$

### 3.2. Extension to arbitrary knot sequences and order of vanishing moments

Observe that (25) can also be derived by applying the formula

$$N'_\ell(x) = N_{\ell-1}(x) - N_{\ell-1}(x-1)$$

to  $N_{2m}(x)$   $m$  times, starting with  $\ell = 2m$  (see, for example, [14, p.86, Theorem 4.3 (vii)]). If the  $n$ -th order vanishing moment in (16) is desired, for any integer  $n \geq 1$ , then the wavelet  $\psi_{m;n}(x)$  of the corresponding interior wavelets  $\psi_{\underline{s};m;n,j}(x)$  in (15) can be formulated as

$$\psi_{\underline{s};m;n,j}(x) = \psi_{m;n}(x - j/2),$$

where

$$\psi_{m;n}(x) := \sum_{k=0}^n (-1)^k \binom{n}{k} N_m(2x - k). \quad (28)$$

The same proof of Theorem 3.1 can be used to yield

$$\psi_{m;n}(x) = N_{m+n}^{(n)}(2x). \quad (29)$$

The importance of the time-domain representation (25), and more generally (29), is that the VM wavelets can be extended to splines on an arbitrary knot sequence

$$\underline{x} : \dots \leq x_j \leq x_{j+1} \leq \dots, \quad (30)$$

with  $x_{j+m} > x_j$  for all  $j \in \mathbb{Z}$ , which includes the knot sequence  $\underline{s}$  in (3), and allows for both bounded intervals and stacked knots, even among the interior knots of the interval. Furthermore, the order  $n$ , of vanishing moments can be arbitrarily chosen, independent of the order  $m$ . In particular, in some applications, it would be preferable to choose  $n > m$ . In addition, the MRA structure considered in the previous subsection is no longer required in our deviation.

**Theorem 3.2** (VM spline-wavelets on arbitrary knot sequences). *Let  $m, n \geq 1$  be arbitrary integers, and  $\underline{x}$  an arbitrary knot sequence as defined in (30). Then the spline basis functions*

$$\psi_{\underline{x},m;n,k}(x) := N_{\underline{x},m+n,k}^{(n)}(x), \quad k \in \mathbb{Z}, \quad (31)$$

*to be called VM wavelets on  $\underline{x}$ , where  $N_{\underline{x},m+n,k}(x)$  is defined in (4), which satisfy the moment conditions*

$$\begin{cases} \int_{-\infty}^{\infty} x^l \psi_{\underline{x},m;n,k}(x) dx = 0, & l = 0, \dots, n-1 \\ \int_{-\infty}^{\infty} x^n \psi_{\underline{x},m;n,k}(x) dx \neq 0. \end{cases} \quad (32)$$

**Remark** Since the knot sequence  $\tilde{s}$  in (17) on the bounded interval  $[0, 2N]$  is a special case of  $\underline{x}$  in (30), the formula (31) can be applied to compute the boundary VM wavelets, such as  $\psi_{\underline{s},4,-2}(x)$  and  $\psi_{\underline{s},4,-1}(x)$ , in the example in Subsection 3.1.

*Proof.* The proof of Theorem 3.2 is a straight-forward application of integration by parts, by using the compact-support property of  $N_{\underline{x},m+n,k}^{(l)}(x)$ ,  $l = 0, \dots, n-1$ , namely,

$$\begin{aligned} & \int_{-\infty}^{\infty} x^l N_{\underline{x},m+n,k}^{(n)} dx \\ &= (-1)^l l! \int_{-\infty}^{\infty} N_{\underline{x},m+n,k}^{(n-l)} dx \\ &= (-1)^l l! [N_{\underline{x},m+n,k}^{(n-l-1)}(x_{m+n+k}) - N_{\underline{x},m+n,k}^{(n-l-1)}(x_k)] = 0 \end{aligned}$$

since  $N_{\underline{x},m+n,k}^{(n-l-1)}(x_{m+n+k}) = N_{\underline{x},m+n,k}^{(n-l-1)}(x_k) = 0$  for all  $l = 0, \dots, n-1$ . On the other hand,

$$\int_{-\infty}^{\infty} x^n N_{\underline{x},m+n,k}^{(n)} dx = (-1)^n n! \int_{x_k}^{x_{m+n+k}} N_{\underline{x},m+n,k} dx \neq 0,$$

since  $N_{\underline{x},m+n,k} > 0$  for  $x_k < x < x_{m+n+k}$  and the assumption that  $x_{m+n+k} > x_k$ .  $\square$

As an immediate consequence of Theorem 3.2, we have the following result.

**Corollary 3.3.** *Let  $m, n \geq 1$  be arbitrary integers. Then the derivative of the VM wavelet  $\psi_{\underline{x},m;n,k}$  on an arbitrary knot sequence  $\underline{x}$ , is given by*

$$\psi'_{\underline{x},m;n,k}(x) = \psi_{\underline{x},m-1;n+1,k}(x). \quad (33)$$

*In particular, for the interior VM wavelets on half-integer knots  $\tilde{\underline{s}}$ ,*

$$\psi'_{\tilde{\underline{s}},m;n,k}(x) = 2\psi_{\tilde{\underline{s}},m-1;n+1,k}(x).$$

*Proof.* Indeed, it follows from (31) in Theorem 3.2 that

$$\psi'_{\underline{x},m;n,k}(x) = N_{\underline{x},m+n,k}^{(n+1)}(x) = N'_{\underline{x},(m-1)+(n+1),k}(x) = \psi_{\underline{x},m-1;n+1,k}(x).$$

□

The importance of the derivative formula (33) is its application to the reassignment technique in the time-frequency analysis which we will discuss in Section 4. In view of the need of taking the  $n^{th}$  order derivatives of the B-splines  $N_{\underline{x},m+n,k}$  to compute the VM wavelet  $\psi_{\underline{x},m;n,k}(x)$ , we end this subsection by recalling the following derivative formula (see, for example, [24])

$$N'_{\underline{x},l,k}(x) = \frac{l-1}{x_{k+l-1} - x_k} N_{\underline{x},l-1,k}(x) - \frac{l-1}{x_{k+l} - x_{k+1}} N_{\underline{x},l-1,k+1}(x),$$

for  $l = m + n, \dots, m + 1$ .

### 3.3. Computation of analytic VM wavelets

In this sub-section, we will focus on the interior wavelets

$$\psi_{m,n}(t) := \psi_{\mathbb{Z},m;n,0}(t)$$

in (31) of Theorem 3.2, with integer knot sequence  $\underline{x} = \mathbb{Z}$  and  $k = 0$ . In signal processing, if  $\psi_{m,n}$  is treated as a signal, then the extension of  $\psi_{m,n}$  to

$$\psi_{m,n}^*(t) := \psi_{m,n}(t) + i(\mathcal{H}\psi_{m,n})(t), \quad (34)$$

called the analytic signal representation allows us to write  $\psi_{m,n}^*(t)$  as an AM-FM signal:

$$\psi_{m,n}^*(t) = A(t)e^{i\phi(t)},$$

with  $A(t) = \sqrt{\psi_{m,n}^2(t) + (\mathcal{H}\psi_{m,n})^2(t)}$  and  $\phi(t) = \tan^{-1}(\mathcal{H}\psi_{m,n}(t)/\psi_{m,n}(t))$ . Here and throughout this paper, given a function  $f$  in the suitable space, the operator  $\mathcal{H}$  denotes the Hilbert transform, defined by

$$(\mathcal{H}f)(t) = p.v. \left\{ \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{f(x)}{t-x} dx \right\}.$$

However, for our application of the VM wavelet as mother wavelets of the CWT, it is necessary to extend  $\psi_{m,n}$  to  $\psi_{m,n}^*$  in order to analyze only the (positive) frequency contents of a given signal  $g(t)$ . Indeed, since

$$\begin{aligned} \widehat{\psi_{m,n}^*}(\omega) &= \widehat{\psi_{m,n}}(\omega) + i(\widehat{\mathcal{H}\psi_{m,n}})(\omega) \\ &= \widehat{\psi_{m,n}}(\omega) + i(-i\text{sgn}(\omega)\widehat{\psi_{m,n}}(\omega)) \\ &= 2\widehat{\psi_{m,n}}(\omega) \end{aligned}$$

for  $\omega > 0$  and  $\widehat{\psi_{m,n}^*}(\omega) = 0$  for  $\omega < 0$ , we have,

$$\begin{aligned} W_g(a, b) &= \langle g, \psi_{m,n}^{*(a,b)} \rangle = \frac{1}{2\pi} \langle \widehat{g}, \widehat{\psi_{m,n}^{*(a,b)}} \rangle \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \widehat{g}(\omega) \overline{\widehat{\psi_{m,n}^*}(\omega)} d\omega \\ &= \frac{1}{2\pi} \int_0^{\infty} \widehat{g}(\omega) \overline{\widehat{\psi_{m,n}^*}(\omega)} d\omega \\ &= \frac{1}{\pi} \int_0^{\infty} \widehat{g}(\omega) \overline{\widehat{\psi_{m,n}}(\omega)} d\omega, \end{aligned}$$

where

$$W_g(a, b) = \langle g, \psi^{(a,b)} \rangle, \quad (35)$$

is the CWT of the function  $g$  in a suitable linear space of functions,  $a > 0$  is the scale,  $b \in \mathbb{R}$  denotes the time position,

$$\psi^{(a,b)}(t) := \frac{1}{a} \psi\left(\frac{t-b}{a}\right), \quad (36)$$

and  $\psi$  is the mother wavelet chosen by the user. To compute  $\mathcal{H}\psi_{m,n}$  in (34), we recall from Theorem 3.2 that

$$\psi_{m,n}(t) = N_{m+n}^{(n)}(t) \quad (37)$$

(for  $\underline{x} = \mathbb{Z}$  and  $k = 0$ ), where  $N_r(t)$  denotes the  $r^{th}$  order cardinal B-spline, defined by the  $r^{th}$ -fold convolution of the first order cardinal B-spline (see (23))

$$N_1(t) = \chi_{[0,1)}(t),$$

so that

$$(\mathcal{H}N_1)(t) = \frac{1}{\pi} \int_0^1 \frac{dx}{t-x} = \ln \left| \frac{t}{t-1} \right|.$$

Also recall the following recurrence formula for computing the  $m^{th}$  order cardinal B-spline  $N_m$  in terms of the characteristic function  $N_1 := \chi_{[0,1)}$ :

$$N_r(t) = \frac{t}{r-1} N_{r-1}(t) + \frac{r-t}{r-1} N_{r-1}(t-1), \quad (38)$$

for  $r = 2, \dots, m$ . One of the important properties of the Hilbert transform for our computation of  $\psi_{m,n}^*$  is that it preserves the recurrence relation (38):

$$(\mathcal{H}N_r)(t) = \frac{t}{r-1} (\mathcal{H}N_{r-1})(t) + \frac{r-t}{r-1} (\mathcal{H}N_{r-1})(t-1), \quad (39)$$

as shown in [11, Theorem 3.2, p.179]. For completeness, we include the derivation of (39) as follows. Let  $I(t) = t$  denote the identity function. Then for  $r \geq 1$ ,

$$\begin{aligned}
(\mathcal{H}(IN_r))(t) &= \frac{1}{\pi} \int_{-\pi}^{\pi} \frac{x N_r(x)}{t-x} dx \\
&= \frac{t}{\pi} \int_{-\infty}^{\infty} \frac{N_r(x)}{t-x} dx - \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{t-x}{t-x} N_r(x) dx \\
&= t(\mathcal{H}N_r)(t) - \frac{1}{\pi} \int_{-\infty}^{\infty} N_r(x) dx \\
&= t(\mathcal{H}N_r)(t) - \frac{1}{\pi}.
\end{aligned}$$

Hence, by applying this identity and the translation-invariance property of the Hilbert transform to the recurrence formula (38), we obtain

$$\begin{aligned}
(\mathcal{H}N_r)(t) &= \frac{1}{r-1} [t(\mathcal{H}N_{r-1})(t) - \frac{1}{\pi}] + \frac{r}{r-1} (\mathcal{H}N_{r-1})(t-1) - \frac{1}{r-1} [t(\mathcal{H}N_{r-1})(t-1) - \frac{1}{\pi}] \\
&= \frac{t}{r-1} (\mathcal{H}N_{r-1})(t) + \frac{r}{r-1} (\mathcal{H}N_{r-1})(t-1) - \frac{t}{r-1} (\mathcal{H}N_{r-1})(t-1),
\end{aligned}$$

which is (39).

We are now ready to derive the following efficient scheme for the computation of  $\mathcal{H}\psi_{m,n}$ , and hence  $\psi_{m,n}^*$ . First, from (37), we observe that

$$\psi_{m,n}(t) = N_{m+n}^{(n)}(t) = \sum_{k=0}^n (-1)^k \binom{n}{k} N_m(t-k).$$

Therefore, by the translation-invariance property of

$$(\mathcal{H}\psi_{m,n})(t) = \sum_{k=0}^m (-1)^k \binom{n}{k} (\mathcal{H}N_m)(t-k), \tag{40}$$

where  $(\mathcal{H}N_m)(t-k)$  can be computed recursively by applying (39), namely,

$$(\mathcal{H}N_r)(t-k) = \frac{t}{r-1} (\mathcal{H}N_{r-1})(t-k) + \frac{r-t}{r-1} (\mathcal{H}N_{r-1})(t-k-1),$$

for  $r = 2, \dots, m$ , with initial function

$$\begin{aligned}
(\mathcal{H}N_2)(t-k) &= t(\mathcal{H}N_1)(t-k) + (2-t)(\mathcal{H}N_1)(t-k-1) \\
&= t \ln \left| \frac{t-k}{t-k-1} \right| + (2-t) \ln \left| \frac{t-k-1}{t-k-2} \right|.
\end{aligned}$$

We remark that although the VM wavelets are compactly supported, their Hilbert transforms  $\mathcal{H}\psi_{m,n}$  do not have finite support. The reason is that

$$(\mathcal{H}N_1)(t) = \ln \left| \frac{t}{t-1} \right| = \ln \left| 1 + \frac{1}{t-1} \right| \sim \left| \frac{1}{t-1} \right|$$

as  $|t| \rightarrow \infty$ , so that

$$\mathcal{H}N_m(t) = O(|t|^{-1})$$

for  $|t| \rightarrow \infty$ , and the decay is not fast. However, since the  $n^{th}$  difference operator acts like the  $n^{th}$  differential operator, it follows from  $\psi_{m,n} = N_{m+n}^{(n)}$  that

$$(\mathcal{H}\psi_{m,n})(t) = O(|t|^{n+1})$$

for  $|t| \rightarrow \infty$ .

We end this section by mentioning that if an VM wavelet  $\psi_{m,n}$  is used as the mother wavelet to compute the CWT of an analytic signal

$$g^*(t) = g(t) + i(\mathcal{H}g)(t),$$

we can avoid the numerical computation of  $\mathcal{H}g$  (which could be expensive and quite inaccurate) by using the analytic VM wavelet  $\psi_{m,n}^*$  to compute the CWT of the given signal  $g(t)$  instead. Indeed, by the anti-symmetry property of the Hilbert transform, we have

$$\begin{aligned} \langle g^*, \psi_{m,n}^{(a,b)} \rangle &= \langle g + i\mathcal{H}g, \psi_{m,n}^{(a,b)} \rangle \\ &= \langle g, \psi_{m,n}^{(a,b)} \rangle + i\langle \mathcal{H}g, \psi_{m,n}^{(a,b)} \rangle \\ &= \langle g, \psi_{m,n}^{(a,b)} \rangle - i\langle g, \mathcal{H}\psi_{m,n}^{(a,b)} \rangle \\ &= \langle g, \psi_{m,n}^{(a,b)} + i\mathcal{H}\psi_{m,n}^{(a,b)} \rangle = \langle g, \psi_{m,n}^{*(a,b)} \rangle \end{aligned}$$

### 3.4. VM wavelets are asymptotically derivatives of the Gaussian function

In this subsection, we first prove that the VM wavelets  $\psi_{m,n}$ , as introduced in (28), are asymptotically the same as the  $n^{th}$  derivative of the Gaussian, after being properly scaled and centered, as  $m+n$  tends to infinity. However, since  $\psi_{m,n}$  is compactly supported, the spectra  $|\hat{\psi}_{m,n}(\omega)|^2$  of  $\psi_{m,n}$  must have side-lobes in nature. The second objective of this subsection is to illustrate the rapid decay of the side-lobes relative to the main-lobes for increasing spline order. We will plot the spectra of  $\psi_m(x) := \psi_{m,m}(x)$  (with  $n = m$ ) for  $m = 3, \dots, 6$  and demonstrate the rapid decrease of the energy of the side-lobe to main-lobe ratios, in terms of SNR (measured in dB).

Let  $g(x)$  be the (normalized) Gaussian function defined by

$$g(x) := \frac{1}{\sqrt{2\pi}} e^{-x^2/2},$$

with Fourier transform given by

$$\hat{g}(\omega) = e^{-\omega^2/2}.$$

In the following, we set  $M = m+n$  and assume that  $m > 1$ . From the definition of  $\psi_{m,n}(x)$  in (29) and following the proof of Theorem 3.1, we have

$$\hat{\psi}_{m,n}(\omega) = \frac{1}{2} \left( \frac{i\omega}{2} \right)^n \hat{N}_M \left( \frac{\omega}{2} \right). \quad (41)$$

We have the following result.

**Theorem 3.4.** For  $M = m + n$ ,

$$\lim_{M \rightarrow \infty} \left[ \left( \frac{M}{12} \right)^{n+1} \psi_{m;n} \left( \sqrt{\frac{M}{48}} (x + 12) \right) - g^{(n)}(x) \right] = 0,$$

where the convergence is both pointwise for all  $x \in \mathbb{R}$  and in  $L^p = L^p(\mathbb{R})$  for  $1 < p < \infty$ .

*Proof.* Taking the Fourier transform of the  $n^{\text{th}}$  derivative of the Gaussian  $g(x)$  and applying the formula (41), we obtain

$$2 \left( \sqrt{\frac{M}{12}} \right)^n \hat{\psi}_{m;n} \left( \frac{4\sqrt{3}\omega}{\sqrt{M}} \right) e^{i\sqrt{3M}\omega} - (i\omega)^n \hat{g}(\omega) = (i\omega)^n \left[ \hat{N}_M \left( \frac{2\sqrt{3}\omega}{\sqrt{M}} \right) e^{i\sqrt{3M}\omega} - \hat{g}(\omega) \right],$$

and by a direct calculation, for a fixed  $\omega$ , we have

$$\hat{N}_M \left( \frac{2\sqrt{3}\omega}{\sqrt{M}} \right) = e^{-i\sqrt{3M}\omega} \left( \frac{\sin(\sqrt{3/M}\omega)}{\sqrt{3/M}\omega} \right)^M.$$

For the Cardinal B-spline  $N_M$ , we have the asymptotic property

$$(i\omega)^n \hat{N}_M \left( \frac{2\sqrt{3}\omega}{\sqrt{M}} \right) e^{i\sqrt{3M}\omega} - (i\omega)^n e^{-\frac{\omega^2}{2}} \rightarrow 0 \quad (42)$$

as  $m \rightarrow \infty$  (see, for example [20, Corollary 1]). Also observe that  $1 \leq p < \infty$ , for all  $M \in \mathbb{N}$ ,

$$(i\omega)^n \hat{N}_M \left( \frac{2\sqrt{3}\omega}{\sqrt{M}} \right) e^{i\sqrt{3M}\omega} - (i\omega)^n \hat{g}(\omega) \in L^p(\mathbb{R}),$$

since  $m > 1$ . In addition, by the same argument as that of [20, Lemma 4.4], for  $1 < p < \infty$ , there exists an  $L^1$  function  $q(\omega)$  so that

$$\left| 2 \left( \sqrt{\frac{M}{12}} \right)^n \hat{\psi}_{m;n} \left( \frac{4\sqrt{3}\omega}{\sqrt{M}} \right) e^{i\sqrt{3M}\omega} - (i\omega)^n \hat{g}(\omega) \right|^p \leq q(\omega). \quad (43)$$

Indeed, for  $|\omega| \leq \frac{\pi}{3} \sqrt{\frac{M}{3}}$ , we have  $\left( \frac{\sin(\sqrt{3/M}\omega)}{\sqrt{3/M}\omega} \right)^M \leq e^{-\omega^2/4}$ ; if  $|\omega| > \frac{\pi}{3} \sqrt{\frac{M}{3}}$ , then  $\left( \frac{\sin(\sqrt{3/M}\omega)}{\sqrt{3/M}\omega} \right)^M \leq \frac{1}{1+|\omega|^M}$ . Hence, the errors in (43) are bounded above by the family of functions  $q_M(\omega)$ , defined by

$$q_M(\omega) := \omega^n \left( e^{-\omega^2/4} + \frac{1}{1+|\omega|^M} \right),$$

which have uniformly bounded  $L^1(\mathbb{R})$  norms, since  $m > 1$ . Thus, by the Lebesgue Dominated Convergence theorem, we may conclude that

$$\lim_{m \rightarrow \infty} \left\| 2 \left( \sqrt{\frac{M}{12}} \right)^n \hat{\psi}_{m;n} \left( \frac{4\sqrt{3}\omega}{\sqrt{M}} \right) e^{i\sqrt{3M}\omega} - (i\omega)^n \hat{g}(\omega) \right\|_{L^p} = 0$$

for all  $1 < p < \infty$ , and hence,

$$\lim_{m \rightarrow \infty} \left\| \left( \frac{M}{12} \right)^{n+1} \psi_{m;n} \left( \sqrt{\frac{M}{48}}(x+12) \right) - g^{(n)}(x) \right\|_{L^2} = 0,$$

where we have applied the Plancherel identity to

$$\begin{aligned} & \mathcal{F}^{-1} \left\{ 2 \left( \sqrt{\frac{M}{12}} \right)^n \hat{\psi}_{m;n} \left( \frac{4\sqrt{3}\omega}{\sqrt{M}} \right) e^{i\sqrt{3M}\omega} - (i\omega)^n \hat{g}(\omega) \right\} (x) \\ &= \left( \frac{M}{12} \right)^{n+1} \psi_{m;n} \left( \sqrt{\frac{M}{48}}(x+12) \right) - g^{(n)}(x), \end{aligned}$$

where  $\mathcal{F}^{-1}$  is the inverse Fourier transform. On the other hand, by a similar argument using the Lebesgue Dominated Convergence theorem, we may conclude that

$$\left( \frac{M}{12} \right)^{n+1} \psi_{m;n} \left( \sqrt{\frac{M}{48}}(x+12) \right) - g^{(n)}(x) \text{ converges to 0 for all } x, \text{ for } m \rightarrow \infty. \quad \square$$

Next, since the spectrum of  $g^{(n)}(x)$  is given by

$$|(i\omega)^n \hat{g}(\omega)|^2 = \omega^{2n} e^{-\omega^2},$$

which may be considered as a band-pass filter with only one main lobe and no side lobes, it is important to investigate the analogous filter property of the VM wavelets  $\psi_{m;n}(x)$ . For convenience, we only consider  $n = m$  and provide the plot of  $|\hat{\psi}_m(\omega)|$  for  $m = 3, \dots, 6$  in Figure 1. The reason for plotting  $|\hat{\psi}_m|$  instead of the spectra  $|\hat{\psi}_m|^2$  is that the side-lobes of  $|\hat{\psi}_m|^2$  are not visible, even for  $m = 3$ . Observe that the side lobes of  $|\hat{\psi}_m(\omega)|$  are visibly much smaller as  $m$  increases from 3 (for quadratic spline) to 6. In fact, for  $m \geq 4$ , the side lobes are practically not visible at all. Hence, we also compute the “side lobe ratio (SLR)”, defined below to compare the energy of the main-lobe (in  $0 \leq \omega \leq \pi$ ) versus the total energy of the side lobes in  $(\pi \leq \omega < \infty)$ , measured in dBs:

$$\text{SLR} := 20 \log_{10} \frac{\|\hat{\psi}_m|_{(0,\pi)}\|_{L^2}}{\|\hat{\psi}_m|_{(\pi,\infty)}\|_{L^2}}. \quad (44)$$

In Figure 2, the plot of the SLR curve in dBs of the main lobe to side lobes (for  $m = 2, \dots, 12$ ) illustrates that the decrease of the side lobe to main lobes ratios to zero is exponentially fast, as  $m$  increases.

#### 4. Real-time time-frequency analysis and time-varying power spectrum

In many applications, the primary purpose of analyzing a time series is to quantify how the dynamics evolves with respect to time. Of course, the interesting dynamical quantities vary from one field of application areas to another. In any case, for a signal with non-stationary oscillatory behavior, it is important to quantify the term “frequency”. For instance, consider the respiratory signal as an example of signals with oscillation. It has been well-known that our breathing rate is not constant and the variability of the



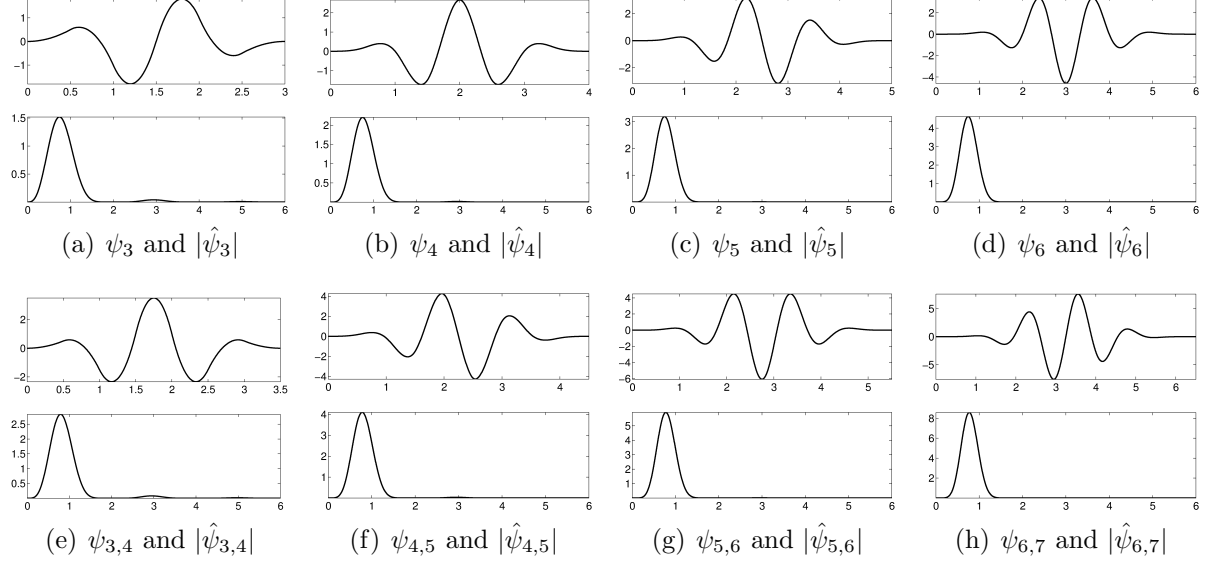


Figure 1: From (a) to (d):  $\psi_3$ ,  $\psi_4$ ,  $\psi_5$  and  $\psi_6$  and the square-root of their power spectra. From (e) to (h):  $\psi_{3,4}$ ,  $\psi_{4,5}$ ,  $\psi_{5,6}$  and  $\psi_{6,7}$  and the square-root their power spectra. The first side lobe in the square-root of the power spectrum of the quadratic VM wavelets  $\psi_3$  and  $\psi_{3,4}$  are clearly visible.

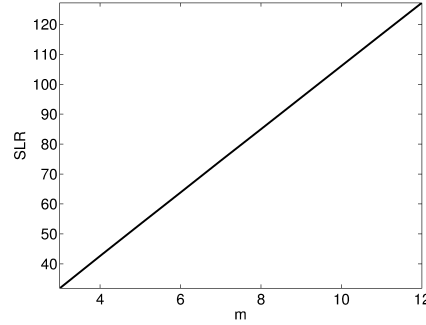


Figure 2: The x-axis is used for the orders  $m$  of the VM wavelets, and the y-axis is for the SLR, measured in dBs.

breathing rate, referred to as breathing rate variability (BRV) (see, for example, [62, 61] and the literature inside), contains information about our health condition. To quantify BRV, and hence the information hidden inside, analyzing the power spectrum (PS) of the respiratory signal, denoted by  $R(t)$ , is the standard yard stick. However, since the notion of frequency is global in nature, the PS is not applicable to momentary oscillatory patterns of  $R(t)$ . In this section, we focus on applying the proposed spline techniques, including the blending operator and VM wavelets, to acquire this dynamical information in the real-time fashion via time-frequency analysis. The result will be applied to the study of the anesthetic depth problem in the next section.

#### 4.1. Adaptive harmonic (or non-harmonic) model

The mathematical model and analysis tools considered have been carefully chosen in the literature, for example, [22, 58, 12], in order to resolve the limitation of the PS. Let

us consider the respiratory signal as an example. As is discussed in [61], our physiological systems closely interact with each other in a complex way. We treat these kinds of signals by a purely phenomenological model, which quantifies the observable features inside the signal. The main characteristic pattern of an oscillatory physiological signal we focus on is its fairly periodic phenomenon. In general, we consider the following *intrinsic-mode typed (IMT) function*:

$$g(t) = \sum_{\ell=1}^K A_{\ell}(t) \cos(2\pi\phi_{\ell}(t)), \quad (45)$$

where  $K$  is finite and for each  $\ell$  the following conditions are satisfied.

$$\begin{cases} A_{\ell} \in C^1(\mathbb{R}) \cap L^{\infty}(\mathbb{R}), & \phi_{\ell} \in C^2(\mathbb{R}), \\ \inf_{t \in \mathbb{R}} A_{\ell}(t) > c_1, & \inf_{t \in \mathbb{R}} \phi'_{\ell}(t) > c_1, \\ \sup_{t \in \mathbb{R}} A_{\ell}(t) \leq c_2, & \sup_{t \in \mathbb{R}} \phi'_{\ell}(t) \leq c_2, \\ |A'_{\ell}(t)| \leq \epsilon \phi'_{\ell}(t), & |\phi''_{\ell}(t)| \leq \epsilon \phi'_{\ell}(t) \quad \text{for all } t \in \mathbb{R}, \end{cases} \quad (46)$$

with  $0 < \epsilon \ll 1$ ,  $0 < c_1 \leq c_2 < \infty$ . In addition,

$$\phi'_{\ell+1}(t) - \phi'_{\ell}(t) > d(\phi'_{\ell+1}(t) + \phi'_{\ell}(t)) \quad (47)$$

for all  $\ell = 1, \dots, K-1$  and  $0 < d < 1$ . We call  $\epsilon, c_1, c_2, d$  *model parameters*. These seemingly complicated mathematical conditions actually have an intuitive interpretation beyond, leading to their nominations.

**Definition** [Phase function and Instantaneous frequency] In the IMT function (45), for each  $\ell$ , the monotonically increasing function,  $\phi_{\ell}(t)$ , is called the *phase function*. The derivative of the phase function,  $\phi'_{\ell}(t)$ , is called the *instantaneous frequency* (IF) of the  $\ell^{th}$  component of the IMT function of  $g(t)$ .

We require the IF to be positive, but usually not constant, allowing it to vary in time, as long as the variations are slight from one “period” to the next for all time  $t$ .

**Definition** [Amplitude modulation] In the IMT function (45), the positive function  $A_{\ell}(t)$  is called the *amplitude modulation* (AM) of  $g(t)$ .

Again, the AM  $A_{\ell}(t)$  is positive, but is allowed to vary. We mention that the motivation and suitability of these definitions have been discussed extensively in [12] and we summarize the result here. Note that we might have infinitely many different ways to represent the same cosine function  $g(t) = \cos(t)$ ; indeed, there exist smooth functions  $\alpha$  and  $\beta$  so that  $g(t) = \cos(t) = (1 + \alpha(t)) \cos(t + \beta(t))$ . However, in general, there is no reason to favor  $\alpha(t) = \beta(t) = 0$ , and hence the definition of AM and IF is problematic. In [12], it is shown that if  $g(t) = A(t) \cos(\phi(t)) = [A(t) + \alpha(t)] \cos(2\pi[\phi(t) + \beta(t)])$  satisfies the IMT condition, then  $|\alpha(t)| \leq C\epsilon$  and  $|\beta'(t)| \leq C\epsilon$ , where  $C$  is a constant depending only on the model parameters  $c_1, c_2, d$ . It is, in this sense, that the IF and AM are well-defined.

**Theorem 4.1.** For an IMT function  $g(t) = \sum_{l=1}^K f_l(t)$ , with  $f_l(t) = A_l(t) \cos(2\pi\phi_l(t))$ , and for all  $t_0 \in \mathbb{R}$  and  $L \gg 1$ , the following “almost orthogonal” relationship holds

$$\left| \int_I \frac{f_k(t)}{\|f_k\|_{L^2(I)}} \frac{f_l(t)}{\|f_l\|_{L^2(I)}} dt \right| \leq \epsilon E + F \quad \text{for all } k \neq l,$$

where  $I := [t_0, t_0 + L]$  and  $E$  and  $F$  are constants to be precisely given in (50) in the derivation. Moreover,  $F$  tends to 0 as  $L \rightarrow \infty$ . In other words, for  $l \neq k$ ,  $f_l$  and  $f_k$  are roughly perpendicular to each other for large values of  $L$ .

*Proof.* Suppose that  $k > l$ . To simplify the notation, we set  $\|\cdot\|_I := \|\cdot\|_{L^2(I)}$ . By a direct computation, we observe that

$$\int_I f_k(t) f_l(t) dt = \frac{1}{2} \int_I A_k(t) A_l(t) [\cos(2\pi(\phi_k(t) - \phi_l(t))) + \cos(2\pi(\phi_k(t) + \phi_l(t)))] dt.$$

Note that by definition,  $\phi'_k(t) - \phi'_l(t) > \left( \left( \frac{1+d}{1-d} \right)^{k-l} - 1 \right) \phi'_l(t) > 0$  for all  $t$ , allowing us to choose  $t_0 \leq t_1 < \dots < t_N \leq t_0 + L$ , so that  $\cos(2\pi(\phi_k(t_l) - \phi_l(t_l))) = 0$  for all  $l = 1, \dots, N$ . Thus, we can divide the integration into pieces which are labeled by (48) and (49) as follows:

$$\begin{aligned} & \int_I A_k(t) A_l(t) [\cos(2\pi(\phi_k(t) - \phi_l(t)))] dt \\ &= \int_{t_0}^{t_1} A_k(t) A_l(t) \cos(2\pi(\phi_k(t) - \phi_l(t))) dt \\ & \quad + \int_{t_N}^{t_0+L} A_k(t) A_l(t) \cos(2\pi(\phi_k(t) - \phi_l(t))) dt \end{aligned} \tag{48}$$

$$+ \sum_{l=1}^{N-1} \int_{t_l}^{t_{l+1}} A_k(t) A_l(t) \cos(2\pi(\phi_k(t) - \phi_l(t))) dt. \tag{49}$$

The estimate for (49) can be achieved simply by applying integration by parts, namely

$$\begin{aligned} & \left| \int_{t_l}^{t_{l+1}} A_k(t) A_l(t) \cos(2\pi(\phi_k(t) - \phi_l(t))) dt \right| \\ &= \left| \int_{t_l}^{t_{l+1}} \partial_t \left( \frac{A_k(t) A_l(t)}{2\pi(\phi'_k(t) - \phi'_l(t))} \right) \sin(2\pi(\phi_k(t) - \phi_l(t))) dt \right| \\ &\leq \epsilon \left| \int_{t_l}^{t_{l+1}} \frac{(\phi'_k(t) A_l(t) + \phi'_l(t) A_k(t))(\phi'_k(t) - \phi'_l(t)) + A_k(t) A_l(t)(\phi'_k(t) + \phi'_l(t))}{2\pi(\phi'_k(t) - \phi'_l(t))^2} dt \right| \\ &\leq \frac{\epsilon}{2\pi(\beta - 1)c_1} \int_{t_l}^{t_{l+1}} (\phi'_k(t) A_l(t) + \phi'_l(t) A_k(t) + \alpha A_k(t) A_l(t)) dt, \end{aligned}$$

where  $\beta := \left( \frac{1+d}{1-d} \right)^{k-l}$  and  $\alpha := \frac{\beta+1}{\beta-1}$ . The last inequality holds since  $\phi'_k(t) + \phi'_l(t) \leq \alpha(\phi'_k(t) - \phi'_l(t))$ . As a result, we have the following bound for (49):

$$\begin{aligned} & \sum_{l=1}^{N-1} \int_{t_l}^{t_{l+1}} A_k(t) A_l(t) \cos(2\pi(\phi_k(t) - \phi_l(t))) dt \\ &\leq \frac{\epsilon}{2\pi(\beta - 1)c_1} \int_{[t_1, t_N]} (\phi'_k(t) A_l(t) + \phi'_l(t) A_k(t) + \alpha A_k(t) A_l(t)) dt. \end{aligned}$$

Due to the lack of cancellation property, we compute an upper bound of the boundary terms (48) simply by  $(L + t_1 - t_N)c_2^2 \leq \frac{2c_2^2}{(\beta-1)c_1}$ , where the last inequality holds due to  $\int_{t_0}^{t_1} \phi'_k(s) - \phi'_l(s)ds = 1 \geq (\beta-1)c_1(t_1 - t_0)$ . Hence, by the Holder's inequality, we obtain

$$\begin{aligned} & \left| \int_I A_k(t) A_l(t) [\cos(2\pi(\phi_k(t) - \phi_l(t)))] dt \right| \\ & \leq \frac{\epsilon}{2\pi(\beta-1)c_1} (\|\phi'_k\|_I \|A_l\|_I + \|\phi'_l\|_I \|A_k\|_I + \alpha \|A_k\|_I \|A_l\|_I) + \frac{2c_2^2}{(\beta-1)c_1}. \end{aligned}$$

Similarly,  $\phi'_k(t) + \phi'_l(t) > 0$  for all  $t$ , so we can choose  $t_0 \leq s_1 < \dots < s_M \leq t_0 + L$  so that  $\cos(2\pi(\phi_k(s_l) + \phi_l(s_l))) = 0$  for all  $l = 1, \dots, M$ . Again, by a similar argument, we have

$$\begin{aligned} & \int_I A_k(t) A_l(t) [\cos(2\pi(\phi_k(t) + \phi_l(t)))] dt \\ & \leq \frac{\epsilon}{2\pi(\beta+1)c_1} (\|\phi'_k\|_I \|A_l\|_I + \|\phi'_l\|_I \|A_k\|_I + \|A_k\|_I \|A_l\|_I) + \frac{2c_2^2}{(\beta+1)c_1}. \end{aligned}$$

In conclusion, we have  $\left| \int_I \frac{f_k(t)}{\|f_k\|_I} \frac{f_l(t)}{\|f_l\|_I} dt \right| \leq E\epsilon + F$ , where

$$\begin{aligned} E &:= \left( \frac{1}{\beta+1} + \frac{1}{\beta-1} \right) \frac{\|\phi'_k\|_I \|A_l\|_I + \|\phi'_l\|_I \|A_k\|_I + \alpha \|A_k\|_I \|A_l\|_I}{4\pi c_1 \|f_k\|_I \|f_l\|_I} \\ F &:= \left( \frac{1}{\beta+1} + \frac{1}{\beta-1} \right) \frac{2c_2^2}{c_1 \|f_k\|_I \|f_l\|_I}. \end{aligned} \tag{50}$$

□

Note that since  $\|A_l\|_I^2 = \int_I f_l^2(t)dt + \int_I |A_l(t)| \sin^2(2\pi\phi_l(t))dt$ , we know that  $\|f_k\|_I \sim \frac{1}{2}\|A_k\|_I$ ; that is,  $\|f_k\|_I$  and  $\|A_k\|_I$  are of the same order no matter how large  $L$  is. Also, by the assumption of  $\phi'_k$  and  $A_k$ , the ratio between  $\|A_k\|_I$  and  $\|\phi'_k\|_I$  is bounded by  $c_2/c_1$  from the above and  $c_1/c_2$  from the below. Thus,  $F$  in (50), which comes from the boundary, decreases to 0 when  $L$  increases, while  $E$  is kept bounded. Also, the  $\beta$  showing in (50) indicates that the farther the instantaneous frequencies of two components are separated, the more these two components are perpendicular to each other.

Next, we shall model the commonly encountered ingredient in signal processing – the trend, of which we consider the following model.

**Definition** [Trend] The trend is a continuous function  $T(t) \in C^0 \cap \mathcal{S}'$  that satisfies

$$\left| \int_0^B \langle T(t), \psi^{(a,b)}(t) \rangle da \right| \ll 1, \quad \left| \int_0^B \langle T(t), \partial_t \psi^{(a,b)}(t) \rangle da \right| \ll 1$$

for all  $b \in \mathbb{R}$  and  $B \gg 1$ , where  $\psi \in \mathcal{S}$  is a mother wavelet,  $\mathcal{S}$  is the Schwartz space and  $\mathcal{S}'$  is the space of tempered distributions. We call  $T(t)$  the *trend associated with  $\psi$  with scale  $B$* .

The trend characterizes the possibly existing non-oscillatory or slowly oscillatory component inside the signal. Indeed, the condition  $|\int_0^B \langle T(t), \psi^{(a,b)}(t) \rangle da| \ll 1$ , for  $B \gg 1$ ,

suggests that at least locally, the trend behaves like a polynomial. Note that polynomials of any finite order and harmonic functions with very slow frequency are special examples of the trend. Thus, we can view the trend as a “time varying DC term” whose local polynomial behavior can be eliminated or minimized by the CWT.

We mention that polynomials of arbitrary degrees are examples of the trend we consider in this paper. To consider another example, let  $\psi(t)$  be any wavelet with compactly supported Fourier transform  $\hat{\psi}(\omega)$  that vanishes on  $0 \leq \omega < 1 - \Delta$ , where  $0 < \Delta < 1$ . Then, for any  $\xi_0 < (1 - \Delta)/B$ , the harmonic function  $\cos(2\pi\xi_0 t)$  can be considered as a trend associated with  $\psi$  with scale  $B$ .

Since all real-world data are contaminated with noise, we consider the following noise model [12].

**Definition** [Noise] Consider a stationary generalized random process  $\Phi$  and a Schwartz function  $\psi$  such that  $\text{var}\Phi(\psi) = 1$  with power spectrum  $d\eta$  of  $\Phi$  satisfies  $\int (1 + |\xi|)^{-2l} d\eta < \infty$  for some  $l > 0$  [12]. Then, the noise model we consider will be  $\sigma(t)\Phi$ , where  $\sigma(t)$  is a slowly varying smooth function so that  $\|\sigma\|_{L^\infty(\mathbb{R})} \ll 1$  and  $\max_{\ell=1,\dots,l} \|\sigma^{(\ell)}\|_{L^\infty} \ll 1$ . We call  $\sigma(t)\Phi$  the noise associated with  $\psi$  with non-stationarity  $\sigma$ .

Note that the continuous-time Gaussian white noise commonly considered in the engineering field, is an example of  $\Phi$ . In addition, the continuous autoregressive and moving average (CARMA) model is also an example of  $\Phi$  [12]. We mention that when  $\sigma$  is a constant,  $\sigma\Phi$  is a stationary noise. The purpose of introducing  $\sigma$  is to capture the possible non-stationary behavior of the noise. We refer the interested reader to [12, Section S.1] for a quick review of the generalized random process.

In summary, in order to capture an observed time series with oscillatory behavior, we may consider the following *adaptive harmonic model associated with a Schwartz function*  $\psi$ :

$$Y = \sum_{\ell=1}^K A_\ell(t) \cos(2\pi\phi_\ell(t)) + T(t) + \sigma(t)\Phi, \quad (51)$$

where  $\sum_{\ell=1}^K A_\ell(t) \cos(2\pi\phi_\ell(t))$  is the IMT function,  $T(t)$  is a trend associated with  $\psi$  with scale  $B$ , and  $\sigma(t)\Phi$  is the noise associated with  $\psi$  with non-stationarity  $\sigma(t)$ . Note that in general  $Y$  is a generalized random process, since by definition,  $\sum_{\ell=1}^K A_\ell(t) \cos(2\pi\phi_\ell(t)) + T(t)$  is a tempered distribution.

Lastly, as is discussed in [59], the cosine function is not always suitable for describing oscillations, for example, the respiratory signal. Indeed, inside a respiratory cycle, the time periods of inhaling and exhaling are different. Thus, the adaptive harmonic model can be generalized to the following *adaptive non-harmonic model*

$$Y(t) = \sum_{\ell=1}^K A_\ell(t) s_\ell(\phi_\ell(t)) + T(t) + \sigma(t)\Phi, \quad (52)$$

where (46) and (47) are satisfied, and each  $s_\ell : [0, 1] \rightarrow \mathbb{R}$  is  $C^{1,\alpha}$ , where  $\alpha > 1/2$  is a 1-periodic function with unit  $L^2$  norm such that  $|\hat{s}_\ell(k)| \leq \delta |\hat{s}_\ell(1)|$  for all  $k \neq 1$ , where  $\delta \geq 0$  is a small parameter, and  $\sum_{n>D} |n\hat{s}_\ell(n)| \leq \theta$  for some small parameter  $\theta \geq 0$  and

$D \in \mathbb{N}$ . The 1-periodic function  $s_\ell(\cdot)$  is called the *wave shape function with dominant ratio  $\delta$ , support  $D$  and accuracy  $\theta$* . We will introduce an algorithm for recovering  $s_\ell$  from the observation later. The interested reader is referred to [59] for further discussions.

#### 4.2. Time-frequency analysis

In general, it is well-known that the PS provided by the Fourier transform is limited to the extraction of the dynamical information of a signal that satisfies the adaptive harmonic (non-harmonic) model. To cope with this limitation, time-frequency analysis was introduced in the literature. Essentially, it is a technique that could be applied to generalize the PS. Intuitively, in order to capture the momentary oscillatory behavior of  $R(t)$ , we may analyze a small piece of the given  $R$  around time  $t_0$  to understand how it oscillates. Mathematically, we realize this idea by working with the short-time Fourier transform (STFT) or the continuous wavelet transform (CWT) of  $R(t)$  (see, for example, [29] for a detailed discussion and other time-frequency analysis techniques). In this paper, we focus ourselves on the CWT, as defined in (35)-(36). We mention that when  $Y$  comes from the adaptive harmonic model,  $W_Y$  is well defined when the mother wavelet  $\psi$  is chosen to be in  $\mathcal{S}$ .

For example, with CWT, we may define the *scalogram* of  $g(t)$  by  $|W_g(a, b)|^2$  [29], which reveals the local oscillatory behavior of  $g(t)$ . We also call  $W_g(a, b)$  the *oscillatory information of the signal associated with scale  $a$  at time  $b$* . Clearly, with the proposed VM wavelet (or any other compactly supported mother wavelet), we may analyze a given signal with CWT and scalogram in real-time, if we are willing to sacrifice the oscillatory information associated with large scales. However, even if the scalogram could be evaluated in real-time, one big issue about the scalogram that remains unsolved is the Heisenberg uncertainty principle – for small scale  $a > 0$ , one cannot expect good temporal resolution. To cope with this difficulty and increase the resolution of the time-frequency analysis, the synchrosqueezed CWT transform (SST) was introduced in [23, 22], and will be discussed in the next section. One reason we focus on the synchrosqueezed CWT is the vanishing moment property of the wavelet transform. Indeed, depending on the chosen mother wavelet, the existing polynomial component in the signal is automatically removed.

#### 4.3. Synchrosqueezed CWT transform and time-varying power spectrum

The synchrosqueezed CWT transform (SST) was first introduced in [23] to analyze voice signals. The initial theoretical understanding of SST is provided in [22], and further studied along with noise and trend in [54, 12]. After its introduction, SST has been applied to a wide range of applications. For example, it has been applied to the medical field for the study of sleep depth via the respiratory signal [12], ventilator weaning prediction problem [61], anesthetic depth estimation via studying the heart rate variability [40, 60], cardiopulmonary coupling analysis [35], etc. We remark that the SST is a special case of the time-frequency reassignment technique [38, 8, 7, 2], and refer the reader to [3] for a recent review article on their relationship and current progress. In addition, SST can also be defined by using the STFT, which was considered and analyzed in [58, 55].

We now summarize the SST theory and algorithm for a given observation  $Y$  for the adaptive harmonic model. We mention that the definition of the Fourier transform in previous papers on SST, for example [54, 12], is different by a  $2\pi$  normalization from what we use in this paper. It is shown that the phase information hidden in CWT

contains information explaining the spread-out phenomenon in the time frequency domain. To remedy it, the CWT coefficients are shifted according to certain reassignment rules determined by the phase information. In this paper, we consider the *reassignment rule* by

$$\Omega_Y(a, b) := \begin{cases} \frac{-i\partial_b W_Y(a, b)}{2\pi W_Y(a, b)} & \text{when } |W_Y(a, b)| > 0 \\ -\infty & \text{when } W_Y(a, b) = 0. \end{cases}$$

where  $\partial_b$  denotes the partial derivative with respect to  $b$ . The synchrosqueezed CWT transform is then achieved by applying this reassignment rule as follows:

$$S_Y^\Gamma(b, \xi) := \int_{\{a: |W_Y(a, b)| > \Gamma\}} W_Y(a, b) \frac{1}{\alpha} h\left(\frac{|\xi - \Omega_Y(a, b)|}{\alpha}\right) a^{-1} da,$$

where  $0 < \alpha \ll 1$  is chosen by the user,  $\Gamma > 0$  is the hard thresholding parameter for reducing the numerical error and noise influence,  $h$  is a smooth function so that  $\frac{1}{\alpha} h(\frac{\cdot}{\alpha}) \rightarrow \delta$ , the Dirac delta distribution, in the weak sense as  $\alpha \rightarrow 0$ , and the scale factor  $a^{-1}$  is used for reconstruction. Note that for each fixed time  $b$ , the CWT coefficient  $W_Y(a, b)$  at scale  $a$  is moved to a new frequency position according to the reassignment rule  $\Omega_Y(a, b)$ . We refer to the square of the modulation of the synchrosqueezed CWT transform, denoted as  $V_Y(b, \xi)$ , as *time-varying power spectrum* (tvPS) of  $Y$ ; that is,

$$V_Y(b, \xi) := |S_Y(b, \xi)|^2.$$

The efficacy of tvPS will be shown in the following numerical subsection. Now we summarize the advantages of SST relevant to our study below:

- (a) SST is robust to several different kinds of noise, which might be slightly non-stationary [12, Theorem 3.1];
- (b) SST is visually informative [12, Theorem 3.1 (ii)]. When the signal is rhythmic, according to the model (45), a dominant curve following the IF can be seen on the tvPS; otherwise the tvPS is blurred ;
- (c) SST is *adaptive* to the data in the sense that its dependence on the chosen window function is weak [12, Theorem 3.1]. (This property is important in the sense that the notion “local” is automatically determined by the data, and hence the model bias can be reduced.)
- (d) SST analysis is local in nature (see the proof of [12, Theorem 3.1]). Indeed, although the conditions about the adaptive harmonic model is stated in the global sense, the analysis simply depends on the local condition of the signal. In this sense, SST, and hence tvPS, can be used to capture the local dynamic behavior of the signal.

These properties have been theoretically justified, and we refer the interested reader to [22, 12] for the proof. The discretized companions of these theorems are discussed in [12], and we refer the reader for the details there. We mention that the condition that the mother wavelet is a Schwartz function is critical in the proof in [12] when the adaptive harmonic model is considered. Indeed, the noise, as a generalized random process, and the trend, as a tempered distribution, cannot be analyzed in general if the Schwartz condition is removed. So, if we use a non-Schwartz mother wavelet, such as the VM wavelets, we have to modify the definition of noise and trend. Since the modification of the proof of the theorem based on these modifications is straightforward but requires additional notations, we will skip the discussion here.

#### 4.4. Synchrosqueezed CWT transform and its numerical limitations

However, from the numerical viewpoint, we have to carefully discuss its implementation limitations. Firstly, the CWT and SST should be implemented in the time domain to avoid undesirable numerical artifacts introduced by the Fourier transform. For example, suppose we have the constant function  $f(t) = c \in \mathbb{R}$ , theoretically  $W_f(a, b) = 0$  for all  $a > 0$  and  $b \in \mathbb{R}$ . However, this is not the case if we only have observation of the signal on a bounded time interval, and this situation is more serious if the observation time is not long enough or if the constant  $c$  is quite large.

Secondly, the differentiation of the CWT in the reassignment rule for the SST algorithm should be exact in order to avoid any possible numerical inaccuracy. Compared with the other mother wavelets with compact supports in the time domain, like the orthogonal Daubechies wavelets [21] and semi-orthogonal spline wavelets [14], the proposed VM wavelets enjoy this property. Indeed, Corollary 3.3 assures that the derivative of any VM wavelet is another VM wavelet. This explicit representation allows us to ease the numerical issue of differentiation.

To sum up, as an application of the VM wavelets, we propose a new numerical implementation of the SST which resolves some of these limitations.

#### 4.5. Causality and real-time implementation of CWT, SST and tvPS

In the above discussion of CWT and SST, the causality issue in data analysis and real-time computing is not considered. Indeed, the mother wavelet such as the Meyer and Morlet wavelets might have infinite support in the time domain so that the future information might contaminate the present information of interest. In addition, real-time implementation is not feasible due to the need of truncation of the chosen mother wavelet in the time-domain. To alleviate these limitations, we might need to choose a mother wavelet with a compact support in time-domain such as the Daubechies wavelets [21] or spline wavelets [14]. However, due to various technical issues, we may only obtain the precise time-frequency information with a lag, depending on the chosen window. Thus, it would be more beneficial to choose a wavelet with the smallest support in the time-domain. On the other hand, we might also want to have the freedom to control the vanishing moment, since in some applications the trend might be modeled by polynomials of lower order. In addition, for the numerical purpose, it is beneficial to have a precise formula for all the kernels involved in the integration. For example, if we can have a precise formula of the derivative of the mother wavelet instead of numerical differentiation, the information we obtain will be more precise.

Thus, based on the above discussion, we propose to use the compactly supported VM wavelet  $\psi_{\underline{x}, m; n, k}(x)$ , where  $\underline{x}$  is the uniform knot sequence and  $m, n \geq 3$ , as the mother wavelet to meet all of the requirements mentioned above. For example, by extending

$$-\partial_b W_g(a, b) = \langle g, \partial_b \psi^{(a, b)} \rangle$$

to

$$\langle g, \partial_b \psi_{\underline{x}, m; n, k}^{(a, b)} \rangle = \langle g, \psi_{\underline{x}, m-1; n+1, k}^{(a, b)} \rangle$$

which follows from Corollary 3.3, where  $\psi_{\underline{x}, m; n, k}$  are VM wavelets on an arbitrary knot sequence  $\underline{x}$ , the reassignment rule  $\Omega(a, b)$  for the SST can be carried out precisely without differentiation.



The most important feature of the proposed solution is actually to settle the lag issue for real-time evaluation. However, near the boundary we may not be able to estimate the CWT coefficient with large scale, or we have to sacrifice the accuracy by either extending the signal or truncating the mother wavelet. With the VM wavelet defined on the nonuniform grid, this issue can be relieved by the extension of the VM wavelet to non-uniform knots, including the derivative of the CWT, by using a boundary wavelet as the mother wavelet. We now introduce the algorithm, as follows.

In this section, the VM wavelets

$$\psi_j := \psi_{\underline{s}, m; m, j} \quad (53)$$

in (15) with  $n = m$  and knot sequence  $\tilde{\mathbf{s}} = \frac{1}{2}\underline{\mathbf{s}}$  (where  $\underline{\mathbf{s}}$  and  $\tilde{\mathbf{s}}$  are defined in (14) and (17) respectively) will be used as the mother wavelets for the wavelet transform of the function  $f(t)$ , which represents the oscillatory pattern, as defined by the adaptive harmonic model (51).

Here, the spline order  $m$  is fixed and the interior VM wavelets  $\psi_j$ , with  $j = 0, \dots, 2(N - m)$ , are precisely translations of the first one

$$\psi_0 = \psi_{\underline{s}, m; m, 0} \quad (54)$$

by  $j/2$ ; that is,

$$\psi_j(x) = \psi(x - j/2), \quad j = 0, \dots, 2(N - m). \quad (55)$$

The omission of various subscripts in the notation (53), (54) and (55) allows us to use the same notation defined in (36) with  $\psi$  as defined in (54), and  $b = j/2$ , for  $j = 0, \dots, 2N - 2m$ . We will denote the “center” interior wavelet by

$$\psi(x) := \psi_0\left(x - \frac{N - m}{2}\right) = \psi_{\underline{s}, m; m, N - m}(x), \quad (56)$$

and apply the scaling and translation notation in (36) to  $\psi(x)$  in (56). Observe that for  $a > 0$ ,  $\psi(x/a)$  is a VM wavelet on the knot sequence.

$$\tilde{\mathbf{s}}_a : 0 < \frac{a}{2} < a < \frac{3a}{2} < \dots < N,$$

and has support inside the time-interval  $[0, N]$ , as long as the scale  $a$  is restricted to  $0 < a < N/m$ . Furthermore, the support of  $\psi(x/a)$  does not overlap with the supports of the boundary wavelets (at both  $t = 0$  and  $t = N$ ) as long as the scale  $a$  is restricted to

$$0 < a \leq \frac{N + 1}{m} - 2. \quad (57)$$

Hence, to allow large values of  $a$  for analyzing low-frequency oscillations, we need a longer time interval  $[0, N]$ .

**Case 1.** ( $1 < a \leq \frac{N+1}{m} - 2$  for low frequency oscillation analysis)

For  $a > 1$ , if  $a$  also satisfies (57), then we may apply the scaling and translation operation to the CWT  $\langle f, \psi^{(a,b)} \rangle$  on the time interval  $[0, N]$  and allow the use of boundary

wavelets to take care of the CWT  $\langle f, \psi_{\underline{s}, m, j} \rangle$  for  $j = -m + 1, \dots, -1$  (for the end-point  $t = 0$  of the time-interval  $[0, N]$ ) and for  $j = 2N - 2m + 1, \dots, 2N - m - 1$  (for the other end-point  $t = N$ ).

**Case 2.** ( $0 < a < 1$  for high-frequency oscillation)

Observe that since the VM wavelets have compact support, scaling by  $a$ ,  $0 < a < 1$ , can also be applied to the interval  $[N - m + \frac{1}{2}, N]$ , on which the  $(m - 1)$  boundary wavelets  $\psi_{\underline{s}, m, j}$ ,  $j = 2N - 2m + 1, \dots, 2N - m - 1$ , live. More precisely, the  $2m - 1$  knots

$$N - m + 1, N - m + 3/2, \dots, N - 1/2 \quad (58)$$

of  $\psi_{\underline{s}, m, j}$ ,  $j = 2N - 2m + 1, \dots, 2N - m + 1$ , in the open interval  $(N - m + 1/2, N)$  can be replaced by  $K$  equally spaced knots, with  $K > 2m - 1$ , namely,

$$y_k := N - m + \frac{1}{2} + k \left( \frac{m - 1/2}{K + 1} \right), \quad k = 1, \dots, K. \quad (59)$$

(Note that  $y_0 = N - m + 1/2$  and  $y_{K+1} = N$  are the two end-points of  $[(N - m) + 1/2, N]$ ) Replacement of the knots in (58) by the knots  $\{y_1, \dots, y_K\}$  in (59) is equivalent to scaling by

$$a := \frac{m - 1/2}{K + 1} = \frac{2m - 1}{2(K + 1)}. \quad (60)$$

We remark that since there are precisely  $m - 1$  boundary wavelets at the end-point  $N$ , the  $K - 2m + 1$  wavelets  $\psi_{\underline{s}_a, m, j}$ ,  $j = 2N - 2m + 1, \dots, 2N - 4m + K + 1$ , with  $a$  given by (60) are interior wavelets with the new equally-spaced knots  $\{y_k\}$  in (59), while the  $m - 1$  wavelets  $\psi_{\underline{s}_a, m, j}$ ,  $j = 2N - 4m + K + 2, \dots, 2N - 3m + K$ , are boundary wavelets. Hence, for  $j = 2N - 2m + 1, \dots, 2N - 4m + K + 1$ , the wavelets  $\psi_{\underline{s}_a, m, j}$  are translations of a single wavelet, namely

$$\psi_{\underline{s}_a, m, j}(x) = \psi_{\underline{s}_a, m, 2N - 2m}(x - (j - (2N - 2m)a)),$$

where  $a$  is given by (60).

#### 4.6. Numerical Implementation and Simulation

The numerical real-time implementation of the SST and tvPS are summarized in Algorithm 2. The Matlab code can be downloaded from <https://sites.google.com/site/hautiengwu/home/download>.

Next we show a numerical simulation of the algorithm. See Figure 4 and Figure 5 for an example that illustrates how well the tvPS captures the dynamical behavior of the signal. In this example we generate the simulated signal in the following way. Simulate the realization of the following random vector of length  $M > 1$

$$\tilde{\phi}_1(l\Delta t) = (W \star K_{\sigma_1})(l\Delta t),$$

---

**Algorithm 2** Summary of the Numerical Implementation

---

- Choose  $m, n \geq 3$ . Take the VM wavelets  $\psi_1 := \psi_{\underline{s}, m; n, 0}$  and  $\psi_2 := \psi_{\underline{s}, m-1; n+1, 0}$ .
- Evaluate the analytic representation of  $\psi_1$  and  $\psi_2$ , denoted by  $\tilde{\psi}_1$  and  $\tilde{\psi}_2$ , by the Hilbert transform (for example, by applying [45, Section 5, p.634–p.635] for the special case of uniform knots);
- Input signal  $Y$  is discretized with sampling period  $\Delta T$ .
- Choose the allowed lag time to be  $L$  seconds so that  $L > (m + n)\Delta T/2$ . Denote  $M = \lfloor L/\Delta T \rfloor$  and  $N = 2M$ .
- Choose  $1/2\Delta T$  (resp.  $1/2L$ ) to be the largest (resp. lowest) frequency of interest. Discretize  $[1/2L, 1/2\Delta T]$  by  $n_\xi$  uniform grids.
- Build up a matrix  $\Psi \in \mathbb{C}^{(N-m-n+1) \times N}$  so that the  $l^{\text{th}}$  entry of the  $i$  row is  $\tilde{\psi}_1^*((m+n)(l-i)/(N-i+1))$ , where  $*$  means the complex conjugate.
- Build up a matrix  $\Lambda \in \mathbb{C}^{(N-m-n+1) \times N}$  so that the  $l^{\text{th}}$  entry of the  $i$  row is  $\tilde{\psi}_2^*((m+n)(l-i)/(N-i+1))$ .
- Start from time 0 and wait up to time  $N\Delta T$ . Set the current time index to be  $I$ .

**while** New input **do**

- $\mathbf{g}_I \leftarrow [g(I-N+1), \dots, g(I)] \in \mathbb{R}^N$ ;
- $W_I \leftarrow \Psi \mathbf{g}_I \in \mathbb{C}^{N-m-n+1}$ ; {Evaluate the CWT at time  $(I-M)\Delta T$ .}
- $Z_I \leftarrow \Lambda \mathbf{g}_I \in \mathbb{C}^{N-m-n+1}$ ; {Evaluate the partial derivative of the CWT with respect to time  $b$  at time  $(I-M)\Delta T$ .}
- $\Omega_I \leftarrow \frac{iZ_I}{2\pi W_I} \in \mathbb{C}^{N-m-n+1}$ , where the division is evaluated entry-wisely; {Evaluate the reassignment rule at time  $(I-M)\Delta T$ .}
- Initiate a zero vector  $S_I \in \mathbb{C}^{n_\xi}$  and evaluate the SST by

**for all**  $j = 1, \dots, N - m - n + 1$  **do**

$$k \leftarrow \text{ROUND} \left[ \frac{\Omega_I(j) - 1/2L}{1/2\Delta T - 1/2L} n_\xi \right]$$

**if**  $1 \leq k \leq n_\xi$  **then**

$$S_I(k) \leftarrow S_I(k) + cW_I(j)a^{-1/2}$$

**end if**

**end for**

- $V_I \leftarrow |S_I|^2 \in \mathbb{R}_+^{n_\xi}$ , where the absolute value and square operators are evaluated entry-wisely. {Evaluate the tvPS at time  $(I-M)\Delta T$ .}

$I \leftarrow I + 1$  { $\Delta T$  second passed.}

**end while**

---

where  $1 \leq l \leq M$ ,  $\Delta t = 1/32$ ,  $W$  is the standard Brownian motion,  $K_{\sigma_1}$  is the Gaussian function of standard deviation  $\sigma_1$  and  $\star$  denotes the convolution operator. The phase function of the first component is defined as

$$\phi_1(K\Delta t) = \Delta t \sum_{l=1}^K 2 \frac{\tilde{\phi}_1(l) + 2 \max_{k=1}^M |\tilde{\phi}_1(k)|}{\max_{l=1}^M \tilde{\phi}_1(l) + 2 \max_{k=1}^M |\tilde{\phi}_1(k)|}.$$

Then, define the AM of the first component by a similar way. For the realization of the random vector of length  $M > 1$

$$\tilde{A}_1(l\Delta t) = (W \star K_{\sigma_2})(l\Delta t),$$

where  $1 \leq l \leq M$ , set

$$A_1(K\Delta t) = \Delta t \sum_{l=1}^K 2 \frac{\tilde{A}_1(l) + 2 \max_{k=1}^M |\tilde{A}_1(k)|}{\max_{l=1}^M \tilde{A}_1(l) + 2 \max_{k=1}^M |\tilde{A}_1(k)|}.$$

We now consider

$$f_1(l\Delta t) := A_1(l\Delta t) \cos(2\pi\phi_1(l\Delta t)) \chi_{1 \leq l \leq 18.75/\Delta t}$$

as the first component. Here,  $1 \leq l \leq M$  and  $\chi$  is the indicator function of  $l$ . The second component, denoted by  $f_2(t)$ , is generated in the same way as that of  $f_1(t)$ , and the trend, denoted by  $T(t)$ , is generated in the same way as that of  $A_1(t)$ , but with non-zero mean removed; that is, the mean of  $T(t)$  is zero. As a result, the final observed signal is  $f(t) = f_1(t) + f_2(t) + T(t)$ . Note that in general, there is no close form expression of  $A_1(t)$  and  $\phi_1(t)$ , and the dynamic of both components can be visually seen from the signal. The signal is shown in Figure 3. To model the noise, we define the signal to noise ratio (SNR) as

$$\text{SNR} := 20 \log_{10} \frac{\text{std}(f_1 + f_2)}{\text{std}(\Phi)}, \quad (61)$$

where  $\text{std}$  means the standard deviation and  $\Phi$  is the noise process. In the following simulations, we choose the VM wavelet  $\psi_{9,9}$  as the mother wavelet with the lag time of 2 seconds. In Figure 4, we show the result of SST when the signal is not contaminated by the noise, and in Figure 5, we show the result of SST when the signal is contaminated by 5dB noise. Note that unlike the PS of  $f(t)$ , it is somewhat obvious that the time-varying dynamics is faithfully reflected in the tvPS. For more robustness results of SST against different noises, we refer the reader of interest to [12].

#### 4.7. Estimation of the wave shape functions

Before closing this section, let us discuss a naive approach to estimate the wave shape functions as introduced in (52) in the adaptive non-harmonic model under the functional regression setup. Consider the wave shape functions  $s_j$ ,  $j = 1, \dots, K$ , with dominant ratio  $\delta > 0$ , support  $D \in \mathbb{N}$  and accuracy  $\theta > 0$  in (52). To simplify our discussion, we assume that  $\theta = 0$ ,  $\sigma = 1$  and  $T = 0$ . Thus, for a function  $f$  that satisfies the adaptive non-harmonic model (52) can be represented by

$$f(t) = \sum_{j=1}^K A_j(t) s_j(\phi_j(t)) = \sum_{j=1}^K A_j(t) \sum_{\ell=1}^D \alpha_{j,\ell} \cos(2\pi\ell\phi_j(t)) + \beta_{j,\ell} \sin(2\pi\ell\phi_j(t)),$$

where  $\alpha_{j,\ell} \in \mathbb{R}$  and  $\beta_{j,\ell} \in \mathbb{R}$  are the Fourier coefficients of the shape function  $s_j$ . We will assume that the observation vector  $\mathbf{Y} \in \mathbb{R}^N$  satisfies

$$\mathbf{Y}(l) = f(l\Delta t) + \Phi_l, \quad (62)$$

where  $l = 1, \dots, N$ ,  $\Delta t > 0$  is the sampling period and  $\Phi$  is a random vector satisfying  $\text{var}(\Phi_l) = 1$  for all  $l$ , which might not be Gaussian and the covariant matrix might not be the identity.

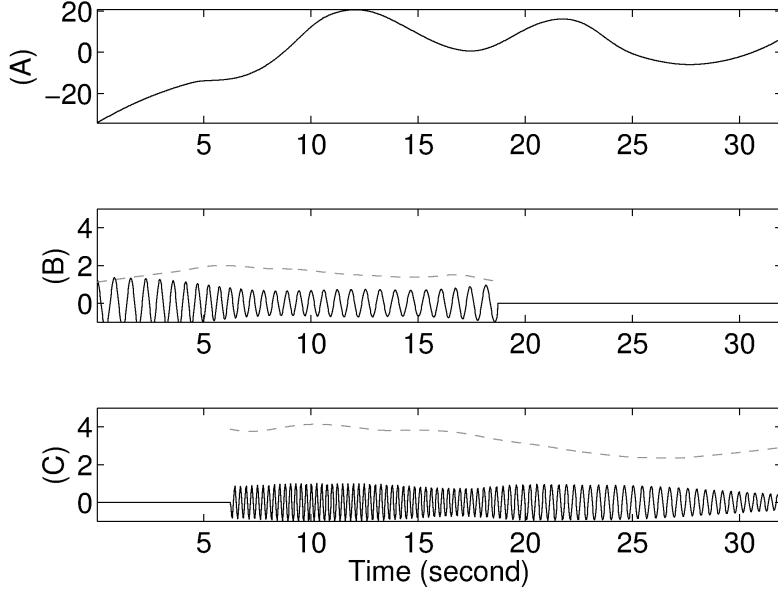


Figure 3: The trend  $T(t)$  is shown in (A). Note that the mean of trend is 0, and its value is  $-20$  in the beginning and about 10 in the end. The first component  $f_1(t)$  and its instantaneous frequency are shown in (B) as the black curve and the gray dashed curve respectively, and the second component  $f_2(t)$  and its instantaneous frequency are shown in (C) as the black curve and the gray dashed curve respectively. Note that  $f_1(t)$  exists up to  $18.75^{th}$  second and  $f_2(t)$  does not exist before the  $6.25^{th}$  second.

The assumption  $\phi'_1(t) < \phi'_k(t)$ , for all  $t \in \mathbb{R}$  and  $k > 1$ , enables us to estimate  $A_1$  and  $\phi_1$  at the sampling points  $l\Delta t$ ,  $l = 1, \dots, N$ , from  $\mathbf{Y}$  via SST with high accuracy; that is, the error is of order  $\epsilon$  (the proof is the same as that in [22, 12] and omitted here). Denote the estimation of  $A_1(t)$  and  $\phi_1(t)$  at the sampling points  $l\Delta t$ ,  $l = 1, \dots, N$ , by  $\tilde{A}_1 \in \mathbb{R}^{1 \times N}$  and  $\tilde{\phi}_1 \in \mathbb{R}^{1 \times N}$ . We construct the estimators  $c_{1,\ell} \in \mathbb{R}^{1 \times N}$  and  $d_{1,\ell} \in \mathbb{R}^{1 \times N}$  by

$$c_{1,\ell}(l) := \tilde{A}_1(l) \cos(2\pi\ell\tilde{\phi}_j(l)), \quad d_{1,\ell}(l) := \tilde{A}_1(l) \sin(2\pi\ell\tilde{\phi}_j(l)),$$

where  $l = 1, \dots, N$  and  $\ell = 1, \dots, D$ . Next, consider the following “functional vectors” with components given by  $c_{1,\ell}$  and  $d_{1,\ell}$ , namely

$$\mathbf{c}_1 = [c_{1,1}^T, \dots, c_{1,D}^T, d_{1,1}^T, \dots, d_{1,D}^T]^T \in \mathbb{R}^{2D \times N}$$

The problem is to evaluate the parameters  $\alpha_{1,\ell}$  and  $\beta_{1,\ell}$ , which are the Fourier coefficients of  $s_1$ , from the functional vectors  $\mathbf{c}_1$ . To facilitate our discussion, let us assume that the estimates  $\tilde{A}_1$  and  $\tilde{\phi}_1$  are precise without error; that is,  $\tilde{A}_1(l) = A_1(l\Delta t)$  and  $\tilde{\phi}_1(l) = \phi_1(l\Delta t)$  for all  $l = 1, \dots, N$ . Then (52) becomes

$$\mathbf{Y}(l) = [\boldsymbol{\gamma}_1 \mathbf{c}_1](l) + \sum_{j=2}^K A_j(l\Delta t) s_j(\phi_j(l\Delta t)) + \Phi_l \quad (63)$$

for all  $l = 1, \dots, N$ , where  $\boldsymbol{\gamma}_1 = [\alpha_{1,1}, \dots, \alpha_{1,D}, \beta_{1,1}, \dots, \beta_{1,D}] \in \mathbb{R}^{1 \times 2D}$ . By Theorem 4.1 which remains valid for the discretization setting, we have

$$\Delta t \mathbf{Y} \mathbf{c}_1^T = \Delta t \boldsymbol{\gamma}_1 \mathbf{c}_1 \mathbf{c}_1^T + \Delta t \Phi \mathbf{c}_1^T + O(\epsilon, (N\Delta t)^{-1}),$$

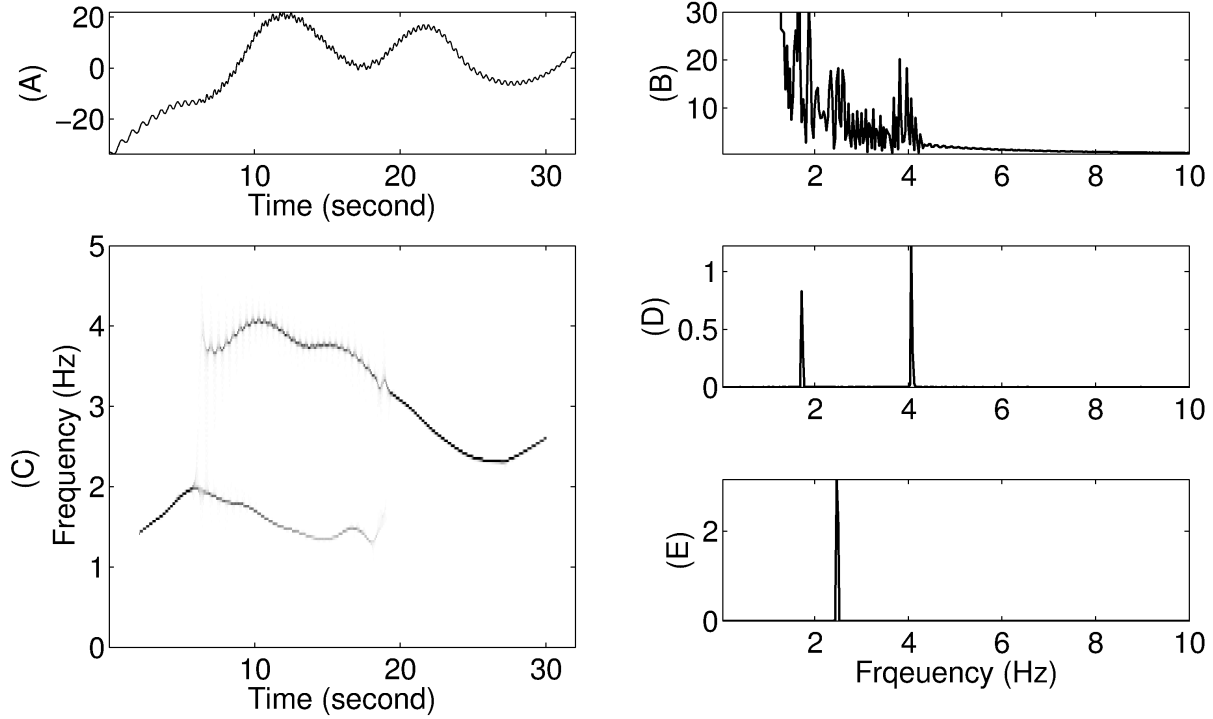


Figure 4: The clean signal  $f(t) := f_1(t) + f_2(t) + T(t)$  is shown in (A). Compared with the oscillatory signal, it is clear that the trend dominates the signal. In (B), we show the power spectrum of  $f(t)$ , where the huge amount of energy in the frequency range  $[0, 1]$  contributed by the trend is discarded to enhance the visualization. Note that the power spectrum is dominated by the trend, and the momentary behavior is difficult to see. The tvPS of  $f(t)$  is shown in (C). Note that the dynamics of  $f_1(t)$  and  $f_2(t)$  can be clearly seen. In (D) and (E), we show the cross section of the tvPS at the 10<sup>th</sup> second and the 24<sup>th</sup> second, respectively. Note that the tvPS is concentrated at the frequency describing the momentary oscillatory behavior of each component, which capture the intuition of “momentary power spectrum”. Also note that the influence of the trend  $T(t)$  in the time-frequency analysis is minimized.

where  $\mathbf{c}_1^T$  is the transpose of  $\mathbf{c}_1$ . Note that since  $\|f_k(t)\|_{L^2([0, N\Delta t])} = \|A_k\|_{L^2([0, N\Delta t])} \gg \epsilon$ , we know the  $2D \times 2D$  matrix  $\mathbf{c}_1 \mathbf{c}_1^T$  is diagonal dominant. Also, by a direct evaluation, we have  $\mathbb{E}(\Delta t \Phi \mathbf{c}_1^T) = 0$  and  $\text{var}(\Delta t \Phi \mathbf{c}_1^T) = O(L\Delta t)$ . Thus we can now estimate the first shape function by using the estimator:

$$\hat{\gamma}_1 := (Y \mathbf{c}_1^T)(\mathbf{c}_1 \mathbf{c}_1^T)^{-1}, \quad (64)$$

where  $\hat{\gamma}_1 = [\hat{\alpha}_{1,1}, \dots, \hat{\alpha}_{1,D}, \hat{\beta}_{1,1}, \dots, \hat{\beta}_{1,D}]^T \in \mathbb{R}^{2D}$ . Indeed, by applying (64), we are able to reconstruct the shape function by

$$\tilde{s}_1(t) := \sum_{\ell=1}^D \hat{\alpha}_{1,\ell} \cos(2\pi \ell t) + \hat{\beta}_{1,\ell} \sin(2\pi \ell t).$$

Then we may iteratively estimate  $A_j(t)$  and  $\phi_j(t)$  for  $j = 2, \dots, K$  and achieve the final result. Note that the success of the regression hinges on Theorem 4.1 and the estimation of  $A_j(t)$  and  $\phi_j(t)$ .

The proposed approach to deal with the adaptive non-harmonic model is based on the functional least-square idea. Since handling the adaptive non-harmonic model is beyond

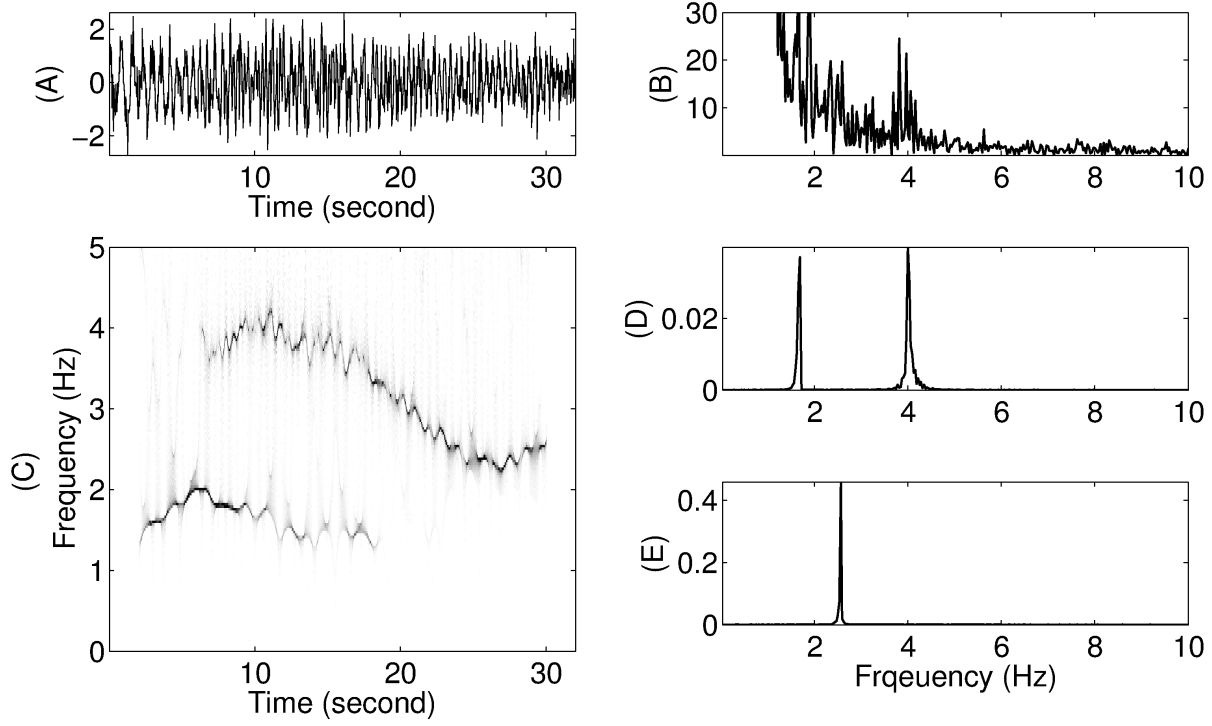


Figure 5: The noisy signal  $f_1(t) + f_2(t) + \Phi$ , where  $\Phi$  is the Gaussian white noise, is shown in (A). For the sake of emphasizing the amount of noise, the trend is not shown. In (B), we show the power spectrum of  $Y = f_1(t) + f_2(t) + T(t) + \Phi$ , and in (C), the tvPS of  $Y$ . (D) and (E) reveal the cross section of the tvPS at the 10<sup>th</sup> second and the 24<sup>th</sup> second, respectively. Note that although the observation is noisy and not too many features can be told from the observation, the dynamics of  $f_1(t)$  and  $f_2(t)$ , like the appearance and disappearance of each component, and the instantaneous frequency, can still be observed. Again, the influence of the trend  $T(t)$  in the time-frequency analysis is minimized.

the scope of this paper, a systematic study, particularly the detailed statistical analysis, will be reported in the near future. Here, we only point out a few possible directions. Firstly, in this subsection we assume the knowledge of  $D$  and  $J$ . However, it might not be the case in practice. Furthermore, for different wave shape functions, the number of non-zero Fourier modes might be less than  $D$ . In this case, we may count on the hypothesis testing to determine the number of significant modes and the number of components. Secondly, the wave shape function model we consider is not the most general possible mode. In some cases we might have a more general, even discontinuous, periodic pattern. In this more general setting, we may apply another approach to recover the shape (see, for example, the diffeomorphism based spectral analysis [63] and singular value decomposition [33]).

## 5. Testbed: Anesthetic Depth Estimation

Respiratory signals contain a wealth of information. A specific information that attracts more and more attention in recent years is the respiratory rate variability (BRV) [62, 61]. In addition to BRV, the analysis of other respiratory patterns shall facilitate the prediction of the patient's prognosis and the choice of the appropriate treatment in some cases [53]. We call these information extracted from the respiratory signal *respiratory dynamical features*. Physiologically, the neural respiratory control is located in the brain

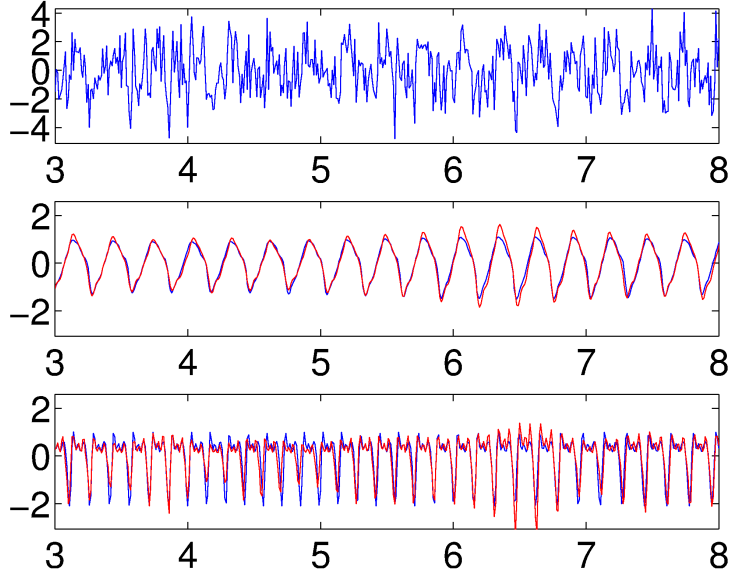


Figure 6: The result of estimating components with non-trivial shape. Top: The signal with two components  $f = f_1 + f_2$ , where  $f_1$  (resp.  $f_2$ ) is non-harmonic oscillatory with shape  $s_1$  (resp.  $s_2$ ), contaminated by the 0-dB Gaussian white noise. Middle: the curve in blue color is the clean signal  $f_1$ , while the red curve is the reconstructed  $f_1$ . Bottom: the curve in blue is the clean signal  $f_2$ , while the curve in red is the reconstructed  $f_2$ .

and comprises two system: the voluntary respiration and the involuntary respiration [46]. In the brainstem, or more specifically the preBötzinger complex [49], the involuntary control center generates more regular oscillatory activity, whereas the voluntary respiration mediated by the control system in the forebrain presents an irregular activity. While competing with each other, the neural controls from these two systems are integrated to control the respiratory motor neuron.

Anesthetics exerting differential effects on the central nerve system, particularly the human brain. It is known that during anesthesia, human respiration is more regular in deeper level of anesthesia, and more irregular in lighter level of anesthesia [37]. This suggests that as the anesthesia depth increases, the anesthetics influences first the forebrain before the brain stem. Despite these early descriptions of the clinical finding, there is no available clinical instrument for quantifying this phenomenon in respiration.

We may model these physiological findings by using the adaptive harmonic model introduced above – when the anesthetic depth is sufficiently deep, the respiratory signal follows the model (45), but the model (45) is not suitable otherwise. In any case, this mathematical model allows us to apply the synchrosqueezing transform to evaluate its tvPS, with which we might be able to evaluate the anesthetic depth evolution during the anesthesia.

In this section, we propose the *blending ECG derived respiration (EDR) algorithm* and apply the proposed real-time SST for estimating respiratory dynamics from the ECG signal in clinical anesthesia. The estimated dynamic is further shown to be closely related to the anesthetic depth.



### 5.1. The blending EDR algorithm

We begin with applying the optimal real-time spline interpolation to obtain the respiratory signal from analyzing the morphology of the ECG signal, which is a commonly encountered non-uniform sampling dataset in the clinics.

The respiratory signal obtained from the ECG measurement is referred to as the *ECG-derived respiration (EDR)* signal. Physiologically, respiration induced ECG distortions occur for two different reasons. Firstly, the respiration-related mechanical changes affect the thoracic electrical impedance, and consequently the cardiac axis rotation occurring during the respiratory cycle has been shown to be the largest factor contributing to the distortion of ECG signals [48]. Secondly, respiration affects the heart rate variability (HRV), thereby causing respiratory sinus arrhythmia [4]. Many algorithms have been developed to estimate the respiration on the basis of the above-mentioned physiological factors. Here we focus on analyzing the ECG morphology variation, which can be viewed as a surrogate of the non-uniform sampling of the respiratory signal. In particular, we consider the one-lead ECG signal, since it is a common setup in several clinical situations.

Now we introduce the EDR algorithm. Given a one-lead ECG signal, for example, lead II or lead III, we first determine the amplitude of the observed R peaks. An R peak is a spiky peak in one heart beat observed in the ECG signal, as illustrated in the middle row of Figure 7 (marked by black crosses). Based on the above mentioned physiological facts, the sampling times of the R peaks and their amplitudes form a non-uniform sampling of the respiratory signal. We comment that the well-known HRV behavior renders these samples irregularly. We then apply the blending operator to interpolate the R peak amplitude data to obtain the respiratory signal in real time. We refer this algorithm as the *blending EDR algorithm*, and we call the result the *blending EDR waveform*.

- (P1) Preprocessing the ECG signal: The median filter is applied to remove the wandering baseline artifact in the ECG signal, for example the lead II ECG signal, which may come from patient movement, dirty lead electrodes and a variety of other things. We choose the moving window of length 100 ms, so that it is longer than the average length of the QRS complex.
- (P2) Time determination of the R peaks (or S peaks when the cardiac axis is deviated): In particular, if the cardiac axis is within the normal range of  $-30^\circ$  to  $90^\circ$  and the ECG signal shows the Rs pattern, we detect the timing of the R peaks; otherwise, we detect the timing of the S peaks with the condition of the rS pattern. If a beat is determined to be premature ventricular complex (PVC), it is deleted.
- (P3) Construction of EDR signal: Build up the EDR signal by applying the blending operator on the amplitudes of the detected R peaks.

The above proposed blending EDR algorithm is illustrated in Figure 7. In summary, we mention that the median filter and the R peak detection can be carried out in real-time, so that by applying the blending operator, the blending EDR waveform can be obtained in real-time. Finally, we remark that traditionally researchers directly apply the cubic spline interpolation of the amplitude of the observed R peaks, which is not real-time.

### 5.2. Nonrhythmic to Rhythmic Ratio

Let us first introduce the necessary terminologies for our purpose. If the signal oscillates in a way that can be quantified by the adaptive harmonic model (52), we say

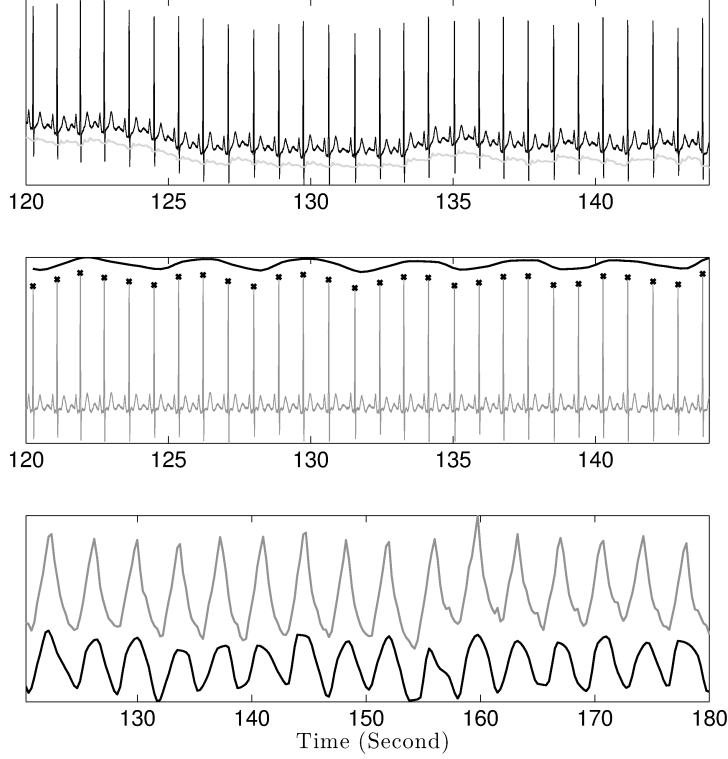


Figure 7: The blending EDR algorithm. In the upper subfigure, the lead II ECG signal is shown as a black curve and the wandering baseline of the lead II ECG signal is plotted as a light gray curve shifted below to enhance the visualization. In the middle subfigure, the median-filtered lead II ECG signal is plotted as a dark gray curve, superimposed with the R peaks marked as black crosses, and the blending EDR waveform is plotted as a black curve shifted up to increase the visualization. In the lower subfigure, the respiratory signal recorded from the chest band is plotted as the dark gray curve while the blending EDR waveform is plotted as the black curve. Note that we can find the oscillatory pattern in the blending EDR waveform, which oscillates similar to the true respiratory signal.

that the signal is *rhythmic*. To be more precise, we consider the following model for the rhythmic blending EDR waveform, namely,

$$R(t) = A(t)s(\phi(t)) + T(t) + \sigma(t)\Phi, \quad (65)$$

that is, with  $K = 1$  in (52). Note that the trend is important in the EDR algorithm. We call the signal which is not rhythmic *non-rhythmic*. A given time series could have a rhythmic behavior at one moment and non-rhythmic at another moment, and this is the dynamics we observe in clinics that we want to quantify. Precisely, as the qualitative description in the literature discussed above, the deeper the anesthetic depth, the more rhythmic the respiration should be, and the proposed index should capture this difference. The proposed index is aimed to quantify this qualitative observation. We now discuss the index to quantify the respiratory dynamic features.

First, discretize the continuous blending EDR waveform  $R(t)$  from the  $0^{th}$  second to the  $L^{th}$  second at  $\eta$ Hz, and evaluate the tvPS of  $R(t)$  with the VM wavelet  $\psi_{m,n}$  and the lag time  $Q$  seconds, which is discretized as a  $n_\xi \times L\eta$  matrix, denoted as  $\mathbf{V}$ , where  $n_\xi$  is the number of discretization of the frequency axis of the tvPS. Here the resolution in the frequency axis is  $\Delta\xi = \frac{\eta}{2n_\xi}$ . With the tvPS, we can define the index

called *Nonrhythmic to rhythmic ratio* (NRR), which was first proposed in [40] to study the dynamic hidden inside HRV. In a nutshell, NRR is defined as a ratio of nonrhythmic component to rhythmic component based on tvPS to differentiate the “nonrhythmic-to-rhythmic” pattern transition. We first define the *dominant frequency*  $f_r$  as the time varying frequency associated with the maximal power on the tvPS, namely

$$f_r = \operatorname{argmax}_{f_r \in Z_{n_\xi}^{L\eta}} \left[ \sum_{m=1}^{L\eta} \log \left( \frac{|\mathbf{V}(f_r(m), m)|}{\sum_{i=1}^{n_\xi} \sum_{j=1}^{L\eta} |\mathbf{V}(j, i)|} \right) - \lambda \sum_{m=2}^{L\eta} |f_r(m) - f_r(m-1)|^2 \right], \quad (66)$$

where  $Z_{n_\xi} = \{1, \dots, n_\xi\}$  and  $\lambda$  is the penalty coefficient determined by the user. Note that when  $R$  satisfies (65), then  $\Delta\xi f_r(l)$ , where  $1 \leq l \leq L\eta$  is simply the instantaneous frequency at time  $l/\eta$  defined in the model [12, Theorem 3.1 (ii)]; otherwise we would obtain the region with highest energy distribution.

Next we define the *rhythmic component power at time  $l/\eta$* , where  $1 \leq l \leq L\eta$ , denoted by  $P_r(l)$ , as the sum of the power inside the bands around  $f_r(l)$  and its multiples on the tvPS, namely

$$P_r(l) := \sum_{k=\lfloor f_r(l)-0.02/\Delta\xi \rfloor}^{\lceil f_r(l)+0.02/\Delta\xi \rceil} \mathbf{V}(l, k),$$

where the width of the band is chosen to be 0.02Hz in an ad hoc way. The *nonrhythmic power at time  $l/\eta$*  is defined as the rhythmic power subtracted from high frequency power of  $R(t)$ ; that is,

$$P_{nr}(l) := \sum_{k=\lceil 0.1/\Delta\xi \rceil}^{n_\xi} \mathbf{V}(l, k) - P_r(l).$$

Finally, NRR at time  $b$  is defined as the ratio of the nonrhythmic component power to the rhythmic component power at time  $l/\eta$ , namely

$$\text{NRR}(l) = \log_{10} \left( \frac{P_{nr}(l)}{P_r(l)} \right).$$

Clearly, when the respiration is rhythmic, NRR is small; while the respiration is non-rhythmic, NRR is large. In this study, we choose  $\eta = 4$ ,  $Q = 45$ ,  $m = n = 11$ ,  $n_\xi = 2000$  and  $\lambda = 0.5$ .

From the data analysis viewpoint, NRR is a time-varying feature extracted from a given oscillatory signal via tvPS, and plays a role of dimensionality reduction. We also mention that although NRR is useful for this anesthetic study, for a different study, we may need a different index derived from the tvPS.

### 5.3. Correlations with Sevoflurane Concentration

To quantify the clinical observation that human respiration is more regular in deeper level of anesthesia, we study the correlation between the NRR index and the inhaled anesthetics (sevoflurane) concentration.

After approval from the Institutional Review Board and obtaining individual written informed consent, we enrolled 31 patients in this study. Physiologic data, EEG data, ECG were recorded continuously and synchronously from standard anesthetic monitoring (HP agilent patient monitor system), Bispectral Index (BIS) monitor (Aspect A-2000

BIS monitor version XP, Host Rev:3.21, smoothing window 15 seconds; Aspect Medical Systems, Nattick, CA, USA) and the ECG recorder (MyECG E3-80; Micro-Star Intl Co., New Taipei City, Taiwan) respectively. The surgery and anesthesia were performed as usual. The inhaled and end-tidal anesthetic gas concentration detected by the gas analyzer on a Datex-Ohmeda S/5 anesthesia machine (GE Health Care, Helsinki, Finland) were also recorded. We focus on the data interval in wakening period from the start of adequate spontaneous breath to return of consciousness when sevoflurane concentration monotonically and continuously decreased.

The NRR is evaluated from the blending EDR waveform derived from the offline lead II ECG data via its tvPS. The inhaled sevoflurane concentration is evaluated by the estimated effect-site (brain) anesthetic gas concentration ( $C_{\text{eff}}$ ), which is derived from the end-tidal anesthetic gas concentration ( $C_{\text{et}}$ ) by the the following pharmacokinetic-pharmacodynamic modeling [52]:

$$\frac{dC_{\text{eff}}}{dt} = K_{\text{e0}}(C_{\text{et}} - C_{\text{eff}}),$$

where the constant  $K_{\text{e0}}$  was defined as 0.20/min for all data [52].

To evaluate the correlation between the NRR index and the inhaled anesthetics (sevoflurane) concentration, we employed the prediction probability ( $P_K$  analysis).  $P_K$  analysis is a standard statistic tool evaluating the performance of anesthetic depth index [51]. Here we briefly summarize the  $P_K$  analysis. Suppose  $x$  is the indicator under analysis, for example, NRR, and  $y$  is the outcome, for example, the concentration of the sevoflurane. Their relationship is described by the rank ordering of pairs of  $(x, y)$ , and there are five possible relationships:

1.  $x$  and  $y$  are *concordant* if pairs of  $(x, y)$  are rank ordered in the same direction;
2.  $x$  and  $y$  are *discordant* if pairs of  $(x, y)$  are rank ordered in the reverse direction

Three different ties in a pair of  $(x, y)$  should be considered:

1. the tie in  $x$ . We call pairs of  $(x, y)$  tie in  $x$  if  $x$ 's are the same while  $y$ 's are different;
2. the tie in  $y$ . We call pairs of  $(x, y)$  tie in  $y$  if  $y$ 's are the same while  $x$ 's are different;
3. the tie in both  $x$  and  $y$ . We call pairs of  $(x, y)$  tie in both  $x$  and  $y$  if  $x$ 's and  $y$ ' are the same.

Among these five relationships, like the usual notion of correlation, in the  $P_K$  analysis the concordance is desirable but not the discordance; unlike the usual notion of correlation, in the  $P_K$  analysis the tie in  $x$  is undesirable but we tolerate the tie in  $y$  and the tie in both  $x$  and  $y$ . To realize this notion, we define  $P_K$  in the following way. Denote  $P_c$ ,  $P_d$  and  $P_{tx}$  the respective probabilities that two pairs of  $(x, y)$  independently drawn from the population with replacement are concordant, discordant and tie in  $x$ . Define

$$P_K = \frac{P_c + \frac{1}{2}P_{tx}}{P_c + P_d + P_{tx}}.$$

We interpret  $P_K$  value in the following way. A value of one means that the indicator always correctly predicts the observed depth of anesthesia, a value of 0.5 means that the indicator predicts no better than 50/50 chance, and a  $P_K$  value less than 0.5 means that the indicator predicts inversely. We mention that the difference between PK analysis and

the commonly used ranking statistics, Spearman correlation, denoted as  $R$ , is that the tie in  $y$  is overlooked in  $P_K$  analysis; it is overlooked since the outcome  $y$  (ex. awake vs. asleep) is usually coarse.

The results were presented as weighted averages according to each patient's data length. As a result, NRR is well correlated with effect-site sevoflurane concentration. Indeed, during spontaneous breathing, the weighted  $P_K$  is  $0.711 \pm 0.021$ . The  $p$  value of this result is  $< 0.0001$  (A  $p$  value less than 0.05 is considered to be significant in this study). Note that this correlation indicates that the respiratory signal contains the dynamical information regarding the anesthesia.

The above preliminary results suggest that the nonrhythmic term and the rhythmic term of the proposed model extract a hidden physiologic information: the relative strength of involuntary and voluntary respiratory controls varies under the influence the dynamic anesthetic effect to the brain. In conclusion, the real-time information of respiratory dynamics revealed in ECG is potential to be a real-time monitoring for clinician to administrate anesthetics. Hence the proposed algorithm could provide an immediate and continuing benefit to the patient.

## 6. Discussion

In addition to the preliminary result about the anesthetic depth evaluation, we mention some potential applications of the proposed algorithms. First, the popular algorithm empirical mode decomposition (EMD) [34] may benefit from it. Indeed, we may replace the cubic spline operator in the sifting process (the key ingredient of the EMD algorithm) by the blending operator so that the real-time implementation of EMD is possible. Second, SST can be applied to analyze the decomposed components (called the intrinsic mode function) to estimate the instantaneous frequency and other information in real-time. Note that if the intrinsic mode functions satisfy the adaptive harmonic model, the result is rigorously justified. A theoretical work toward understanding the sifting process, its real-time implementation, and a survey of its feasibility to data analysis will be reported in the future paper. In particular, an exact and efficient computational scheme for the analytic VM wavelets

$$\psi_{\underline{x},m;n,j}^* = \psi_{\underline{x},m;n,j} + i\mathcal{H}\psi_{\underline{x},m;n,j},$$

with arbitrary knot sequence  $\underline{x}$  on bounded intervals will be derived (based on the Hilbert B-spline in [45]), to facilitate the wavelet analysis of the sifting process

$$g(t) = \sum_{k=1}^K g_k(t) + r(t)$$

for an arbitrarily given signal  $g$  on a bounded interval, without the need of numerical integration (in computing the analytic intrinsic mode signal components  $g_k^* = g_k + i\mathcal{H}g_k$ ). The reason, as already discussed at the end of Section 3.3, is that

$$\langle g_k^*, \psi_{\underline{x},m;n,j} \rangle = \langle g_k, \psi_{\underline{x},m;n,j}^* \rangle.$$

In Section 4.3, we already mentioned some applications of SST in the medical field. Here we briefly mention that SST itself can be applied to other fields as well. Examples

include: adaptive seasonality estimation in the high frequency financial data [56], non-stationary dynamics analysis in the financial system [30], gearbox fault diagnosis [39], paleoclimatic data analysis [54], seasonal behavior of diseases in epidemiology [12], etc. It is worth noting that while combined with the blending operator and VM wavelets, SST may be applied to the problems where online monitoring of the system dynamics is important.

Note that in Algorithm 2, we need the analytic representation of the VM wavelet  $\psi_1$  by applying the Hilbert transform. We show the necessity of the Hilbert transform by illustrating with the following example. Suppose  $f(t) = \cos(2\pi t)$ ; that is, we have a pure-tone harmonic function with frequency 1, or an IMT function with a constant IF 1 and constant AM 1. By a direct calculation, we have, for any mother wavelet  $\psi_1$ ,

$$W_f(a, b) = \langle f, \psi_1^{(a,b)} \rangle = \frac{1}{2}\sqrt{a}\hat{\psi}_1(a)e^{i2\pi b} + \frac{1}{2}\sqrt{a}\hat{\psi}_1(-a)e^{-i2\pi b},$$

and hence

$$-i\partial_b W_f(a, b) = \pi\sqrt{a}\hat{\psi}_1(a)e^{i2\pi b} - \pi\sqrt{a}\hat{\psi}_1(-a)e^{-i2\pi b}.$$

It is clear that  $\Omega_f(a, b)$  does not provide any information about the frequency 1. However, if we consider  $\psi_1^*$  as the analytic representation of  $\psi_1$ , we see that

$$W_f(a, b) = \langle f, \psi_1^{*(a,b)} \rangle = \frac{1}{2}\sqrt{a}\hat{\psi}_1^*(a)e^{i2\pi b}, \quad -i\partial_b W_f(a, b) = \pi\sqrt{a}\hat{\psi}_1^*(a)e^{i2\pi b},$$

which indicates that  $\Omega_f(a, b) = 1$  when  $|W_f| > 0$ .

There are several open problems left unanswered from the viewpoint of data analysis. For example, how is the number of components determined from a given observation, under the assumption that the clean signal satisfies the adaptive harmonic (non-harmonic) model? A possible solution is proposed in [12] by using the hypothesis test. However, a systematic study is lacking. Even if we know the oracle about the number of the components, what is the proper notion of statistical accuracy of extracting the information when noise exist, and how to optimize the extraction accuracy? Also, when the noise is non-stationary, the above problems become more challenging. One possible approach to the above problems is to model the AM and IF functions in terms of splines so that the adaptive harmonic (non-harmonic) model becomes parametric. Under this situation, the above problems could become more trackable and several traditional statistical techniques may be applied.

## Acknowledgment

The research of Charles K. Chui was supported by the U.S. Army Research Office under Grant #W911NF-11-1-0426. Hau-Tieng Wu acknowledges support by AFOSR grant FA9550-09-1-0643. The authors are also indebted to Maryke van der Walt for carrying out the computations of the cubic VM wavelet example on a bounded interval in Section 3.1.

## References

- [1] Y. Ait-Sahalia and P. A. Mykland. The effects of random and discrete sampling when estimating continuous time diffusions. *Econometrica*, 71(2):483–549, 2003.

- [2] F. Auger, E. Chassande-Mottin, and P. Flandrin. Making reassignment adjustable: The levenberg-marquardt approach. In *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, pages 3889–3892, March 2012.
- [3] F. Auger, P. Flandrin, Y.-T. Lin, S. McLaughlin, S. Meignen, T. Oberlin, and H.-T. Wu. Recent Advances in Time-Frequency Reassignment and Synchrosqueezing. *IEEE Trans. Signal Process.*, 30:32–41, 2013.
- [4] R. Bailón, L. Sornmo, and P. Laguna. Ecg derived respiratory frequency estimation. In G. D. Clifford, F. Azuaje, and P. E. McSharry, editors, *Advanced Methods and Tools for ECG Data Analysis*, pages 215–243. London: Artech House, 2006.
- [5] J. J Benedetto. Irregular sampling and frames. In C.K. Chui, editor, *wavelets: A Tutorial in Theory and Applications*, pages 445–507. Academic Press, Boston, 1992.
- [6] I. Bilinskis and A. K. Mikelson. *Randomized Signal Processing*. Prentice Hall, 1992.
- [7] E. Chassande-Mottin, F. Auger, and P. Flandrin. Time-frequency/time-scale reassignment. In *Wavelets and signal processing*, Appl. Numer. Harmon. Anal., pages 233–267. Birkhäuser Boston, Boston, MA, 2003.
- [8] E. Chassande-Mottin, I. Daubechies, F. Auger, and P. Flandrin. Differential reassignment. *IEEE Signal Process. Lett*, 4(10):293–294, Oct 1997.
- [9] C.-Y. Chen and P. P. Vaidyanathan. Mimo radar space – time adaptive processing using prolate spheroidal wave functions. *Signal Processing, IEEE Transactions on*, 56(2):623–635, Feb 2008.
- [10] G. Chen, C. K. Chui, and M. Lai. Construction of real-time spline quasi-interpolation schemes. *Approx. Th. and its Appl.*, 4:61–75, 1988.
- [11] Q. Chen, N. Huang, S. Riemenschneider, and Y. Xu. A b-spline approach for empirical mode decomposition. *Adv. Comput. Math.*, 24:171–195, 2006.
- [12] Y.-C. Chen, M.-Y. Cheng, and H.-T. Wu. Nonparametric and adaptive modeling of dynamic seasonality and trend with heteroscedastic and dependent errors. *J. Roy. Stat. Soc. B*, 76, 2014. in press.
- [13] C. K. Chui. *Multivariate splines*. SIAM, 1988.
- [14] C. K. Chui. *An Introduction to Wavelets*. Academic Press, 1992.
- [15] C. K. Chui. *Wavelets: A Mathematical Tool for Signal Analysis*. SIAM, 1997.
- [16] C. K. Chui and G. Chen. *Kalman filtering: with real-time applications*. Springer, N.Y., 2009.
- [17] C. K. Chui and J. M. De Villiers. Applications of optimally local interpolation to interpolatory approximates and compactly supported wavelets. *Math. Comp.*, 213:99–114, 1996.

- [18] C. K. Chui and H. Diamond. A general framework to local interpolation. *Num. Math.*, 58:569 – 581, 1991.
- [19] C. K. Chui and M. J. Lai. Computation of box splines and b-splines on triangulations of nonuniform rectangular partitions. *Approx. Theory Appl.*, 3:37–62, 1987.
- [20] C. K. Chui and J. Z. Wang. A study of asymptotically optimal time-frequency localization by scaling functions and wavelets. *Ann. of Num. Math.*, 4:193–216, 1997.
- [21] I. Daubechies. *Ten Lectures on Wavelets*. Philadelphia: Society for Industrial and Applied Mathematics, 1992.
- [22] I. Daubechies, J. Lu, and H.-T. Wu. Synchrosqueezed Wavelet Transforms: an empirical mode decomposition-like tool. *Appl. Comput. Harmon. Anal.*, pages 243–261, 2010.
- [23] I. Daubechies and S. Maes. A nonlinear squeezing of the continuous wavelet transform based on auditory nerve models. In A. Aldroubi and M. Unser. Boca Raton, editors, *Wavelets in Medicine and Biology*, pages 527–546. CRC-Press, 1996.
- [24] C. de Boor. *A practical guide to splines*. Springer-Verlag, 1978.
- [25] C. de Boor and G. Fix. Spline approximation by quasi-interpolants. *Journal of Approximation Theory*, 8:96–110, 1973.
- [26] J. de De Villiers. *Mathematics of Approximation*. Atlantis Press, 2012.
- [27] Y. C. Eldar. Sampling with arbitrary sampling and reconstruction spaces and oblique dual frame vectors. *Journal of Fourier Analysis and Applications*, 9(1):77–96, 2003.
- [28] J. Fan and I. Gijbels. *Local Polynomial Modelling and Its Applications*. Chapman and Hall/CRC, 1996.
- [29] P. Flandrin. *Time-frequency/time-scale Analysis, Wavelet Analysis and Its Applications*. Academic Press Inc., 1999.
- [30] S. Guharay, G. Thakur, F. Goodman, S. Rosen, and D. Houser. Analysis of non-stationary dynamics in the financial system. *Economics Letters*, 121:454–457, 2013.
- [31] F. Gunnarsson, F. Gustafsson, and F. Gunnarsson. Frequency analysis using non-uniform sampling with application to active queue management. In *Acoustics, Speech, and Signal Processing, 2004. Proceedings. (ICASSP '04). IEEE International Conference on*, volume 2, pages ii–581–4 vol.2, May 2004.
- [32] L. G. Gurin, B. T. Polyak, and É. V. Raik. The method of projections for finding the common point of convex sets. *Zh. Vychisl. Mat. Mat. Fiz.*, 7(6):1211–1228, 1967.
- [33] T. Y. Hou, Z. Shi, and P. Tavallali. Sparse time frequency representations and dynamical systems. *CoRR*, abs/1312.0202, 2013.



- [34] N. E. Huang, Z. Shen, S. R. Long, M.C. Wu, H.H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. Lond. A*, 454(1971):903–995, 1998.
- [35] D. Iatsenko, A. Bernjak, T. Stankovski, Y. Shiogai, P.J. Owen-Lynch, P. B. M. Clarkson, P. V. E. McClintock, and A. Stefanovska. Evolution of cardiorespiratory interactions with age Evolution of cardiorespiratory interactions with age. *Phil. Trans. R. Soc. A*, 371(20110622):1–18, 2013.
- [36] K. Jacobsson, L.L.H. Andrew, A.K. Tang, S.H. Low, and H. Hjalmarrsson. An improved link model for window flow control and its application to fast tcp. *Automatic Control, IEEE Transactions on*, 54(3):551–564, March 2009.
- [37] M. Kato, T. Komatsu, T. Kimura, F. Sugiyama, K. Nakashima, and Y. Shimada. Spectral analysis of heart rate variability during isoflurane anesthesia. *Anesthesiology*, 77:669–674, 1992.
- [38] K. Kodera, R. Gendrin, and C. Villedary. Analysis of time-varying signals with small bt values. *IEEE Trans. Acoust., Speech, Signal Processing*, 26(1):64 – 76, feb 1978.
- [39] C. Li and M. Liang. Timefrequency signal analysis for gearbox fault diagnosis using a generalized synchrosqueezing transform. *Mechanical Systems and Signal Processing*, 26:205–217, 2012.
- [40] Y.-T. Lin, H.-T. Wu, J. Tsao, H.-W. Yien, and S.-S. Hseu. Time-varying spectral analysis revealing differential effects of sevoflurane anaesthesia: non-rhythmic-to-rhythmic ratio. *Acta Anaesthesiologica Scandinavica*, 58:157–167, 2014.
- [41] N. R. Lomb. Least-squares frequency analysis of unequally spaced data. *Astrophysics and Space Science*, 39:447–462, February 1976.
- [42] T. Lyche and L.L. Schumaker. Local spline approximation. *Journal of Approximation Theory*, 15:294–325, 1975.
- [43] M. J. Marsden. An identity for spline functions with applications to variation-diminishing spline approximation. *J. Approx. Theory*, 3:7–49, 1970.
- [44] F. Marvasti. *Nonuniform Sampling: Theory and Practice*. Springer, 2001.
- [45] C. A. Micchelli, Y. Xu, and B. Yu. On computing with the hilbert spline transform. *Adv. Comput. Math.*, 38:623–646, 2013.
- [46] R.A. Mitchell and A.J. Berger. Neural regulation of respiration. *Am Rev Respir Dis*, 111:206–224, 1975.
- [47] G. B. Moody. Spectral analysis of heart rate without resampling. In *Computers in Cardiology 1993, Proceedings.*, pages 715–718, Sep 1993.
- [48] G. B. Moody, R. G. Mark, A. Zoccola, and S. Mantero. Derivation of respiratory signals from multi-lead ecgs. *Comput. Cardiol.*, 12:113–116, 1985.

- [49] J.C. Rekling and J.L. Feldman. PreBötzinger complex and pacemaker neurons: hypothesized site and kernel for respiratory rhythm generation. *Annu Rev Physiol*, 60:385–405, 1998.
- [50] L. L. Schumaker. *Spline functions: Basic theory*. Wiley-Interscience, 1981.
- [51] W.D. Smith, R.C. Dutton, and N.T. Smith. Measuring the performance of anesthetic depth indicators. *Anesthesiology*, 84:38–51, 1996.
- [52] M. Soehle, R.K. Ellerkmann, M. Grube, M. Kuech, S. Wirz, A. Hoeft, and J. Bruhn. Comparison between Bispectral Index and Patient State Index as Measures of the Electroencephalographic Effects of Sevoflurane. *Anesthesiology*, 109:799–805, 2008.
- [53] V. K. Somers, D. P. White, R. Amin, W. T. Abraham, F. Costa, A. Culebras, S. Daniels, J. S. Floras, C. E. Hunt, L. J. Olson, T. G. Pickering, R. Russell, M. Woo, and T. Young. Sleep apnea and cardiovascular disease: An american heart association/american college of cardiology foundation scientific statement from the american heart association council for high blood pressure research professional education committee, council on clinical cardiology, stroke council, and council on cardiovascular nursing in collaboration with the national heart, lung, and blood institute national center on sleep disorders research (national institutes of health). *J. Am. Coll. Cardiol.*, 52(8):686–717, 2008.
- [54] G. Thakur, E. Brevdo, N. S. Fuckar, and H.-T. Wu. The synchrosqueezing algorithm for time-varying spectral analysis: robustness properties and new paleoclimate applications. *Signal Processing*, 93:1079–1094, 2013.
- [55] G. Thakur and H.-T. Wu. Synchrosqueezing-based recovery of instantaneous frequency from nonuniform samples. *SIAM J. Math. Anal.*, 43:2078–2095, 2011.
- [56] T. Vatter, H.-T. Wu, V. Chavez-Demoulin, and B. Yu. Non-parametric estimation of intraday spot volatility: disentangling instantaneous trend and seasonality. *SSRN e-prints*, 2013. 2330159.
- [57] R. Vio, T. Strohmmer, and W. Wamsteker. On the reconstruction of irregularly sampled time series. *Publications of the Astronomical Society of the Pacific*, 112(767):74–90, 2000.
- [58] H.-T. Wu. *Adaptive Analysis of Complex Data Sets*. PhD thesis, Princeton University, 2011.
- [59] H.-T. Wu. Instantaneous frequency and wave shape functions (I). *Appl. Comput. Harmon. Anal.*, 35:181–199, 2013.
- [60] H.-T. Wu, Y.-H. Chan, Y.-T. Lin, and Y.-H. Yeh. Using synchrosqueezing transform to discover breathing dynamics from ecg signals. *Appl. Comput. Harmon. Anal.*, 36:354–359, 2014.
- [61] H.-T. Wu, S.-S. Hseu, M.-Y. Bien, Y. R. Kou, and I. Daubechies. Evaluating physiological dynamics via synchrosqueezing: Prediction of ventilator weaning. *IEEE Transactions on Biomedical Engineering*, 61:736–744, 2013.

- [62] M. Wysocki, C. Cracco, A. Teixeira, A. Mercat, J. Diehl, Y. Lefort, J Derenne, and T Similowski. Reduced breathing variability as a predictor of unsuccessful patient separation from mechanical ventilation. *Crit. Care Med.*, 34:2076–2083, 2006.
- [63] H. Yang. Synchrosqueezed Wave Packet Transforms and Diffeomorphism Based Spectral Analysis for 1D General Mode Decompositions. *ArXiv e-prints 1311.4655 [math.NA]*, 2013.